

=> d his ful

(FILE 'HOME' ENTERED AT 10:31:06 ON 31 AUG 2006)

FILE 'REGISTRY' ENTERED AT 10:32:46 ON 31 AUG 2006

L1 3 SEA ABB=ON PLU=ON VETWFLRHP|IETWFLRHP|RETWFLRHP|VESWFLRNP/SQSP

FILE 'CAPLUS' ENTERED AT 10:34:23 ON 31 AUG 2006

L2 2 SEA ABB=ON PLU=ON L1

=> fil reg

FILE 'REGISTRY' ENTERED AT 10:35:17 ON 31 AUG 2006  
 USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
 PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
 COPYRIGHT (C) 2006 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file  
 provided by InfoChem.

STRUCTURE FILE UPDATES: 30 AUG 2006 HIGHEST RN 905475-39-0  
 DICTIONARY FILE UPDATES: 30 AUG 2006 HIGHEST RN 905475-39-0

New CAS Information Use Policies, enter HELP USAGETERMS for details.

TSCA INFORMATION NOW CURRENT THROUGH June 30, 2006

Please note that search-term pricing does apply when  
 conducting SmartSELECT searches.

REGISTRY includes numerically searchable data for experimental and  
 predicted properties as well as tags indicating availability of  
 experimental property data in the original document. For information  
 on property searching in REGISTRY, refer to:

<http://www.cas.org/ONLINE/UG/regprops.html>

=> d que 11

L1 3 SEA FILE=REGISTRY ABB=ON PLU=ON VETWFLRHP|IETWFLRHP|RETWFLRHP  
 |VESWFLRNP/SQSP

=> d sqide3 11 1-3

L1 ANSWER 1 OF 3 REGISTRY COPYRIGHT 2006 ACS on STN  
 RN 849707-86-4 REGISTRY  
 CN L-Proline, L-seryl-L-valyl-L-alanyl-L-leucyl-L-valyl-L-prolyl-L-histidyl-L-  
 valylglycyl-L-methionylglycyl-L-leucyl-L- $\alpha$ -glutamyl-L-threonyl-L-  
 arginyl-L-threonyl-L- $\alpha$ -glutamyl-L-threonyl-L-tryptophyl-L-methionyl-  
 L-seryl-L-seryl-L- $\alpha$ -glutamylglycyl-L-alanyl-L-tryptophyl-L-lysyl-L-  
 histidyl-L-valyl-L-glutaminyl-L-arginyl-L-isoleucyl-L- $\alpha$ -glutamyl-L-  
 threonyl-L-tryptophyl-L-phenylalanyl-L-leucyl-L-arginyl-L-histidyl- (9CI)  
 (CA INDEX NAME)

OTHER NAMES:

CN 46: PN: US20050080231 FIGURE: 6 claimed sequence  
 FS PROTEIN SEQUENCE  
 SQL 40

PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
Not Given	US2005080231
	claimed
	FIGURE 6

SEQ3 1 Ser-Val-Ala-Leu-Val-Pro-His-Val-Gly-Met-  
 11 Gly-Leu-Glu-Thr-Arg-Thr-Glu-Thr-Trp-Met-  
 21 Ser-Ser-Glu-Gly-Ala-Trp-Lys-His-Val-Gln-



31 Arg-Ile-Glu-Thr-Trp-Phe-Leu-Arg-His-Pro

=== ===

HITS AT: 32-40

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

DT.CA CAPLUS document type: Patent

RL.P Roles from patents: BIOL (Biological study); PRP (Properties)

1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

L1 ANSWER 2 OF 3 REGISTRY COPYRIGHT 2006 ACS on STN

RN 819099-23-5 REGISTRY

CN L-Proline, L-leucyl-L-alanyl-L-methionyl-L- $\alpha$ -glutamyl-L- $\alpha$ -glutamyl-L-leucyl-L-tyrosyl-L-arginyl-L-seryl-L-valyl-L-alanyl-L-leucyl-L-valyl-L-prolyl-L-histidyl-L-valylglycyl-L-methionylglycyl-L-leucyl-L- $\alpha$ -glutamyl-L-threonyl-L-arginyl-L-threonyl-L- $\alpha$ -glutamyl-L-threonyl-L-tryptophyl-L-methionyl-L-seryl-L-seryl-L- $\alpha$ -glutamylglycyl-L-alanyl-L-tryptophyl-L-lysyl-L-histidyl-L-valyl-L-glutaminyl-L-arginyl-L-isoleucyl-L- $\alpha$ -glutamyl-L-threonyl-L-tryptophyl-L-phenylalanyl-L-leucyl-L-arginyl-L-histidyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 34: PN: US20040266987 SEQID: 35 unclaimed protein

FS PROTEIN SEQUENCE

SQL 48

PATENT ANNOTATIONS (PNTE):

Sequence | Patent

Source | Reference

=====+=====

Not Given | US2004266987

| unclaimed

| SEQID 35

SEQ3 1 Leu-Ala-Met-Glu-Glu-Leu-Tyr-Arg-Ser-Val-

11 Ala-Leu-Val-Pro-His-Val-Gly-Met-Gly-Leu-

21 Glu-Thr-Arg-Thr-Glu-Thr-Trp-Met-Ser-Ser-

31 Glu-Gly-Ala-Trp-Lys-His-Val-Gln-Arg-Ile-

===

41 Glu-Thr-Trp-Phe-Leu-Arg-His-Pro

=== ===

HITS AT: 40-48

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

DT.CA CAPLUS document type: Patent

RL.P Roles from patents: PRP (Properties)

1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

L1 ANSWER 3 OF 3 REGISTRY COPYRIGHT 2006 ACS on STN

RN 819099-19-9 REGISTRY

CN L-Proline, L-seryl-L-valyl-L-alanyl-L-leucyl-L-valyl-L-prolyl-L-histidyl-L-valylglycyl-L-methionylglycyl-L-leucyl-L- $\alpha$ -glutamyl-L-threonyl-L-arginyl-L-threonyl-L- $\alpha$ -glutamyl-L-threonyl-L-tryptophyl-L-methionyl-

L-seryl-L-seryl-L- $\alpha$ -glutamylglycyl-L-alanyl-L-tryptophyl-L-lysyl-L-histidyl-L-glutamyl-L-arginyl-L-isoleucyl-L- $\alpha$ -glutamyl-L-threonyl-L-tryptophyl-L-phenylalanyl-L-leucyl-L-arginyl-L-histidyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 29: PN: US20040266987 SEQID: 29 unclaimed protein  
FS PROTEIN SEQUENCE  
SQL 39

PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
=====+=====	
Not Given	US2004266987
	unclaimed
	SEQID 29

SEQ3 1 Ser-Val-Ala-Leu-Val-Pro-His-Val-Gly-Met-  
11 Gly-Leu-Glu-Thr-Arg-Thr-Glu-Thr-Trp-Met-  
21 Ser-Ser-Glu-Gly-Ala-Trp-Lys-His-Gln-Arg-  
31 Ile-Glu-Thr-Trp-Phe-Leu-Arg-His-Pro  
=== === === === === === === ===

HITS AT: 31-39

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

DT.CA CAPLUS document type: Patent

RL.P Roles from patents: PRP (Properties)

1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

=> fil caplus

FILE 'CAPLUS' ENTERED AT 10:36:20 ON 31 AUG 2006

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2006 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 31 Aug 2006 VOL 145 ISS 10

FILE LAST UPDATED: 30 Aug 2006 (20060830/ED)

Effective October 17, 2005, revised CAS Information Use Policies apply. They are available for your review at:

<http://www.cas.org/infopolicy.html>

'OBI' IS DEFAULT SEARCH FIELD FOR 'CAPLUS' FILE

=> d que 12

L1 3 SEA FILE=REGISTRY ABB=ON PLU=ON VETWFLRHP|IETWFLRHP|RETWFLRHP  
|VESWFLRNP/SQSP  
L2 2 SEA FILE=CAPLUS ABB=ON PLU=ON L1

=> d .ca 12 1-3

L2 ANSWER 1 OF 2 CAPLUS COPYRIGHT 2006 ACS on STN  
ACCESSION NUMBER: 2005:325740 CAPLUS  
DOCUMENT NUMBER: 142:385982  
TITLE: Small peptides having apoptotic activities and their applications  
INVENTOR(S): Despres, Philippe; Catteau, Adeline  
PATENT ASSIGNEE(S): Institut Pasteur, Fr.  
SOURCE: U.S. Pat. Appl. Publ., 43 pp., Cont.-in-part of U.S. Ser. No. 311,213.  
CODEN: USXXCO  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 2  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2005080231	A1	20050414	US 2003-608147	20030630
WO 2001096376	A2	20011220	WO 2001-IB1570	20010618
WO 2001096376	A3	20030313		
WO 2001096376	C2	20031023		
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 2004101862	A1	20040527	US 2003-311213	20030519
US 2004049016	A1	20040311	US 2003-634895	20030806
PRIORITY APPLN. INFO.:				
			US 2000-212129P	P 20000616
			WO 2001-IB1570	W 20010618
			US 2003-311213	A2 20030519
			US 2001-881710	A3 20010618

OTHER SOURCE(S): MARPAT 142:385982

ED Entered STN: 15 Apr 2005

AB The present invention relates to nine residue peptides (M32-40) from flavivirus M ectodomain able to modulate specifically the apoptotic activity of diverse flavivirus, to pharmaceutical composition comprising the same and their use for the treatment and/or the prevention of flavivirus-linked infections and cancers. Dengue virus M ectodomain peptide fusion protein containing the M precursor translocation signal sequence induced apoptosis in mouse neuroblastoma Neuro 2a and human hepatoma HepG2 cancer cells.

IC ICM C07K007-08

ICS C07K007-06

INCL 530329000; 530330000

CC 1-6 (Pharmacology)

Section cross-reference(s): 10

IT 849707-84-2 849707-85-3 849707-86-4

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
(Biological study)

(M protein fragment, apoptosis induction in relation to; small  
apoptotic peptides from flavivirus M ectodomain for treating and  
preventing flavivirus-linked infections and cancers)

L2 ANSWER 2 OF 2 CAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2005:2227 CAPLUS

DOCUMENT NUMBER: 142:86611

TITLE: Attenuated flavivirus strains containing a mutated  
M-ectodomain and their applications

INVENTOR(S): Despres, Philippe; Catteau, Adeline

PATENT ASSIGNEE(S): Institut Pasteur, Fr.

SOURCE: U.S. Pat. Appl. Publ., 30 pp.

CODEN: USXXCO

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2004266987	A1	20041230	US 2003-608029	20030630
			US 2003-608029	20030630

PRIORITY APPLN. INFO.:

ED Entered STN: 31 Dec 2004

AB The present invention relates to nine residue peptides (ApoptoM) from  
flavivirus M ectodomain able to modulate specifically the apoptotic  
activity of diverse flavivirus, to pharmaceutical composition comprising the  
same and their use for the treatment and/or the prevention of  
flavivirus-linked infections and cancers.

IC ICM C12Q001-70

ICS A61K039-12; A61K039-193; C07K002-00; C07K004-00; C07K005-00;  
C07K007-00; C07K014-00; C07K016-00; C07K017-00; A61K038-00;  
C07K001-00

INCL 530300000; 530350000; 424204100; 424218100

CC 1-5 (Pharmacology)

Section cross-reference(s): 6, 10

IT 819099-13-3 819099-14-4 819099-15-5 819099-19-9  
819099-23-5

RL: PRP (Properties)

(unclaimed protein sequence; attenuated flavivirus strains containing a  
mutated M-ectodomain and their applications)

=>

=> d his ful

FILE 'CAPLUS' ENTERED AT 10:37:19 ON 31 AUG 2006

E DENGUE/CT

E E3+ALL

L1 1995 SEA ABB=ON PLU=ON DENGUE/OBI  
 L2 19 SEA ABB=ON PLU=ON ( M/OBI (L) ECTODOMAIN/OBI OR ECTO  
 DOMAIN/OBI)  
 L3 51 SEA ABB=ON PLU=ON ( M (3A) ECTODOMAIN OR ECTO DOMAIN)/AB  
 L4 737 SEA ABB=ON PLU=ON ECTODOMAIN/OBI OR ECTO DOMAIN/OBI  
 L5 5 SEA ABB=ON PLU=ON L4 AND L1

FILE 'MEDLINE, BIOSIS, EMBASE, WPIX' ENTERED AT 10:40:44 ON 31 AUG 2006

L6 14584 SEA ABB=ON PLU=ON DENGUE  
 L7 6991 SEA ABB=ON PLU=ON ECTODOMAIN OR ECTO DOMAIN  
 L8 22 SEA ABB=ON PLU=ON L6 AND L7

FILE 'CAPLUS, MEDLINE, BIOSIS, EMBASE, WPIX' ENTERED AT 10:41:25 ON 31  
 AUG 2006

L9 12 DUP REM L5 L8 (15 DUPLICATES REMOVED)  
 ANSWERS '1-5' FROM FILE CAPLUS  
 ANSWERS '6-9' FROM FILE MEDLINE  
 ANSWERS '10-12' FROM FILE WPIX  
 L10 265 SEA ABB=ON PLU=ON DESPRES P?/AU  
 E LA GARENNE COLOMBES/AU  
 E COLOMBES LA G/AU  
 E LA GARENNE/AU  
 L11 65 SEA ABB=ON PLU=ON CATTEAU A?/AU  
 L12 308 SEA ABB=ON PLU=ON (L10 OR L11)  
 L13 85 SEA ABB=ON PLU=ON L12 AND DENGUE  
 L14 74 SEA ABB=ON PLU=ON L13 NOT L8  
 L15 0 SEA ABB=ON PLU=ON L14 AND (ECTODOMAIN OR ECTO DOMAIN)  
 L16 2 SEA ABB=ON PLU=ON L14 AND PEPTIDE#  
 L17 26 SEA ABB=ON PLU=ON L14 AND SEQUENCE#  
 D TI 1-10  
 L18 26 SEA ABB=ON PLU=ON L16 OR L17  
 L19 14 DUP REM L18 (12 DUPLICATES REMOVED)  
 ANSWERS '1-6' FROM FILE CAPLUS  
 ANSWERS '7-9' FROM FILE MEDLINE  
 ANSWER '10' FROM FILE BIOSIS  
 ANSWER '11' FROM FILE EMBASE  
 ANSWERS '12-14' FROM FILE WPIX

=> fil caplus medline biosis embase wpix  
FILE 'CAPLUS' ENTERED AT 10:47:18 ON 31 AUG 2006  
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
COPYRIGHT (C) 2006 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'MEDLINE' ENTERED AT 10:47:18 ON 31 AUG 2006

FILE 'BIOSIS' ENTERED AT 10:47:18 ON 31 AUG 2006  
Copyright (c) 2006 The Thomson Corporation

FILE 'EMBASE' ENTERED AT 10:47:18 ON 31 AUG 2006  
Copyright (c) 2006 Elsevier B.V. All rights reserved.

FILE 'WPIX' ENTERED AT 10:47:18 ON 31 AUG 2006  
COPYRIGHT (C) 2006 THE THOMSON CORPORATION

=> d que 19  
L1 1995 SEA FILE=CAPLUS ABB=ON PLU=ON DENGUE/OBI  
L4 737 SEA FILE=CAPLUS ABB=ON PLU=ON ECTODOMAIN/OBI OR ECTO  
DOMAIN/OBI  
L5 5 SEA FILE=CAPLUS ABB=ON PLU=ON L4 AND L1  
L6 14584 SEA DENGUE  
L7 6991 SEA ECTODOMAIN OR ECTO DOMAIN  
L8 22 SEA L6 AND L7  
L9 12 DUP REM L5 L8 (15 DUPLICATES REMOVED)

=> d que 119  
L6 14584 SEA DENGUE  
L7 6991 SEA ECTODOMAIN OR ECTO DOMAIN  
L8 22 SEA L6 AND L7  
L10 265 SEA DESPRES P?/AU  
L11 65 SEA CATTEAU A?/AU  
L12 308 SEA (L10 OR L11)  
L13 85 SEA L12 AND DENGUE  
L14 74 SEA L13 NOT L8  
L16 2 SEA L14 AND PEPTIDE#  
L17 26 SEA L14 AND SEQUENCE#  
L18 26 SEA L16 OR L17  
L19 14 DUP REM L18 (12 DUPLICATES REMOVED)

=> d .ca 19 1-5; d ibib ab ct 19 6-12; d ibib ab 119 1-14

L9 ANSWER 1 OF 12 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 1  
ACCESSION NUMBER: 2005:1350338 CAPLUS  
DOCUMENT NUMBER: 144:86572  
TITLE: Antibodies specific to West Nile virus E proteins for  
diagnosis, prophylaxis and treatment of flavivirus  
infection  
INVENTOR(S): Fikrig, Erol; Gould, Hannah; Koski, Raymond A.;  
Ledizet, Michel; Marasco, Wayne A.  
PATENT ASSIGNEE(S): USA  
SOURCE: PCT Int. Appl., 111 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2005123774	A2	20051229	WO 2005-US22188	20050615
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW			
RW:	BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			

PRIORITY APPLN. INFO.:  
 US 2004-580248P P 20040615  
 US 2004-613369P P 20040927  
 US 2005-646839P P 20050124

ED Entered STN: 30 Dec 2005

AB The present invention relates to anti-West Nile virus E protein (WNE) antibodies, including human antibodies, and antigen-binding portions thereof. In particular, the invention relates to such antibodies and portions that prevent, inhibit, or treat a flavivirus infection, including a West Nile Virus infection. The invention also relates to antibodies that are chimeric, bispecific, derivatized, single chain antibodies or that are portions of fusion proteins. The invention also relates to isolated heavy and light chain Igs derived from human anti-WNE antibodies and nucleic acid mols. encoding such Igs. The present invention also relates to methods of making human anti-WNE antibodies, compns. comprising these antibodies and methods of using the antibodies and compns. for diagnosis, prophylaxis and treatment. The invention also provides gene therapy methods using nucleic acid mols. encoding the heavy and/or light Ig mols. that comprise the human anti-WNE antibodies. The invention also relates to transgenic animals or plants comprising nucleic acid mols. of the present invention.

IC ICM C07K016-00

CC 15-3 (Immunochemistry)

IT Protein motifs

(ectodomain I and II; human antibodies specific to West Nile virus E proteins for diagnosis, prophylaxis and treatment of flavivirus infection)

IT Adoptive immunotherapy

B cell (lymphocyte)

Blood serum

Dengue virus

Dengue virus 1

Dengue virus 2

Dengue virus 3

Dengue virus 4

Dissociation constant

Drug delivery systems

Epitopes

Flavivirus

Genetic vectors

Human

Immunoassay

Japanese encephalitis virus

Kunjin virus

Molecular cloning  
Murray Valley encephalitis virus  
Phage display library  
Protein sequences  
St. Louis encephalitis virus  
West Nile virus  
cDNA sequences  
(human antibodies specific to West Nile virus E proteins for diagnosis,  
prophylaxis and treatment of flavivirus infection)

L9 ANSWER 2 OF 12 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 2  
ACCESSION NUMBER: 2005:2227 CAPLUS  
DOCUMENT NUMBER: 142:86611  
TITLE: Attenuated flavivirus strains containing a mutated M-  
**ectodomain** and their applications  
INVENTOR(S): Despres, Philippe; Catteau, Adeline  
PATENT ASSIGNEE(S): Institut Pasteur, Fr.  
SOURCE: U.S. Pat. Appl. Publ., 30 pp.  
CODEN: USXXCO  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2004266987	A1	20041230	US 2003-608029	20030630
PRIORITY APPLN. INFO.:				
			US 2003-608029	20030630

ED Entered STN: 31 Dec 2004  
AB The present invention relates to nine residue peptides (ApoptoM) from  
flavivirus M ectodomain able to modulate specifically the apoptotic  
activity of diverse flavivirus, to pharmaceutical composition comprising the  
same and their use for the treatment and/or the prevention of  
flavivirus-linked infections and cancers.  
IC ICM C12Q001-70  
ICS A61K039-12; A61K039-193; C07K002-00; C07K004-00; C07K005-00;  
C07K007-00; C07K014-00; C07K016-00; C07K017-00; A61K038-00;  
C07K001-00  
INCL 530300000; 530350000; 424204100; 424218100  
CC 1-5 (Pharmacology)  
Section cross-reference(s): 6, 10  
ST flavivirus M protein **ectodomain** peptide apoptosis induction  
vaccine  
IT Proteins  
RL: BPN (Biosynthetic preparation); DGN (Diagnostic use); THU (Therapeutic  
use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(M (matrix); attenuated flavivirus strains containing a mutated m-  
**ectodomain** and their applications)  
IT Antiviral agents  
Dengue virus 1  
Dengue virus 2  
Flavivirus  
Immunoassay  
Japanese encephalitis virus  
Protein sequences  
Vaccines  
Viral RNA sequences  
Yellow fever virus  
(attenuated flavivirus strains containing a mutated m-**ectodomain**  
and their applications)



- IT Drug delivery systems  
(carriers; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Protein motifs  
(**ectodomain**, pro-apoptosis peptide derived from; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Human  
(flavivirus infection detection in; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Diagnosis  
(mol.; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Antibodies and Immunoglobulins  
RL: ARG (Analytical reagent use); DGN (Diagnostic use); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)  
(monoclonal, to **dengue** virus 2 M protein **ectodomain** ; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Immobilization, molecular or cellular  
(of **dengue** virus 2 M protein **ectodomain** derived peptide; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Fusion proteins (chimeric proteins)  
RL: BPN (Biosynthetic preparation); DGN (Diagnostic use); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(of **dengue** virus 2 M protein **ectodomain** derived peptide; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Plasmid vectors  
(pC95-114-EGFPM1-40(I136F)DEN-2; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Peptides, biological studies  
RL: BSU (Biological study, unclassified); DGN (Diagnostic use); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
(pro-apoptosis, derived from **dengue** virus 2 M protein **ectodomain**; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Flavivirus  
(recombinant, containing mutated protein **ectodomain**; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Mutagenesis  
(site-directed, of **dengue** virus 2 M protein **ectodomain**; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT Antibodies and Immunoglobulins  
RL: ARG (Analytical reagent use); DGN (Diagnostic use); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)  
(to **dengue** virus 2 M protein **ectodomain**; attenuated flavivirus strains containing a mutated m-**ectodomain** and their applications)
- IT 819074-11-8  
RL: PRP (Properties)  
(Unclaimed; attenuated flavivirus strains containing a mutated M-**ectodomain** and their applications)
- IT 174515-88-9    212568-00-8    212622-13-4    212900-31-7    212900-32-8  
212900-33-9    255698-48-7    485261-07-2    486165-36-0    486165-37-1  
486165-38-2    486165-43-9    486196-81-0    486196-82-1    489821-07-0  
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

(Biological study)  
 (amino acid sequence; attenuated flavivirus strains containing a mutated m-  
**ectodomain** and their applications)

IT 819098-43-6D, derivs. claimed  
 RL: BSU (Biological study, unclassified); DGN (Diagnostic use); PRP  
 (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (dengue virus 2 M protein **ectodomain** derived  
 peptide; attenuated flavivirus strains containing a mutated m-  
**ectodomain** and their applications)

IT 140977-36-2 164447-31-8, GenBank U21055 165915-82-2, GenBank U17066  
 165915-83-3, GenBank U17067 210508-08-0, GenBank AF052437 210508-09-1,  
 GenBank AF052438 210508-10-4, GenBank AF052439 210508-24-0, GenBank  
 AF052444 210508-25-1, GenBank AF052445 210508-26-2, GenBank AF052446  
 221425-89-4, GenBank AF094612 384469-39-0 385281-78-7 385281-79-8  
 385281-80-1  
 RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
 (Biological study)  
 (nucleotide sequence; attenuated flavivirus strains containing a mutated m-  
**ectodomain** and their applications)

IT 819098-92-5 819098-93-6 819098-94-7 819098-95-8 819098-96-9  
 819098-97-0 819098-98-1 819098-99-2 819099-00-8 819099-01-9  
 819099-02-0 819099-03-1 819099-04-2 819099-05-3 819099-06-4  
 819099-07-5 819099-08-6 819099-09-7 819099-10-0 819099-11-1  
 819099-12-2 819099-16-6 819099-17-7 819099-18-8 819099-20-2  
 819099-21-3 819099-22-4 819099-24-6  
 RL: PRP (Properties)  
 (unclaimed nucleotide sequence; attenuated flavivirus strains containing a  
 mutated M-**ectodomain** and their applications)

IT 819099-13-3 819099-14-4 819099-15-5 819099-19-9 819099-23-5  
 RL: PRP (Properties)  
 (unclaimed protein sequence; attenuated flavivirus strains containing a  
 mutated M-**ectodomain** and their applications)

IT 819074-12-9  
 RL: PRP (Properties)  
 (unclaimed sequence; attenuated flavivirus strains containing a mutated M-  
**ectodomain** and their applications)

L9 ANSWER 3 OF 12 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 3

ACCESSION NUMBER: 2004:51428 CAPLUS

DOCUMENT NUMBER: 140:177012

TITLE: Structure of the dengue virus envelope  
 protein after membrane fusion

AUTHOR(S): Modis, Yorgo; Ogata, Steven; Clements, David;  
 Harrison, Stephen C.

CORPORATE SOURCE: Howard Hughes Medical Institute, Children's Hospital  
 and Harvard Medical School, Boston, MA, 02115, USA

SOURCE: Nature (London, United Kingdom) (2004), 427(6972),  
 313-319

CODEN: NATUAS; ISSN: 0028-0836

PUBLISHER: Nature Publishing Group

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 21 Jan 2004

AB Dengue virus enters a host cell when the viral envelope glycoprotein, E,  
 binds to a receptor and responds by conformational rearrangement to the  
 reduced pH of an endosome. The conformational change induces fusion of  
 viral and host-cell membranes. A three-dimensional structure of the soluble  
 E ectodomain (sE) in its trimeric, postfusion state reveals striking  
 differences from the dimeric, prefusion form. The elongated trimer bears  
 three fusion loops' at one end, to insert into the host-cell membrane.

Their structure allows us to model directly how these fusion loops interact with a lipid bilayer. The protein folds back on itself, directing its carboxy terminus towards the fusion loops. We propose a fusion mechanism driven by essentially irreversible conformational changes in E and facilitated by fusion-loop insertion into the outer bilayer leaflet. Specific features of the folded-back structure suggest strategies for inhibiting flavivirus entry.

CC 6-3 (General Biochemistry)  
 Section cross-reference(s): 75  
 ST **dengue** virus envelope protein E conformation cell membrane bilayer  
 IT Envelope proteins  
 RL: BSU (Biological study, unclassified); PEP (Physical, engineering or chemical process); PRP (Properties); PYP (Physical process); BIOL (Biological study); PROC (Process)  
 (E; conformational changes of **dengue** virus envelope protein E enhances its fusion to cell membrane through fusion-loop insertion)  
 IT Membrane, biological  
 (bilayer, lipid; conformational changes of **dengue** virus envelope protein E enhances its fusion to cell membrane through fusion-loop insertion)  
 IT Cell membrane  
 Conformational transition  
 (conformational changes of **dengue** virus envelope protein E enhances its fusion to cell membrane through fusion-loop insertion)  
 IT Protein motifs  
 (ectodomain; conformational changes of **dengue** virus envelope protein E enhances its fusion to cell membrane through fusion-loop insertion)  
 IT Conformation  
 (loop, protein, three fusion loops; conformational changes of **dengue** virus envelope protein E enhances its fusion to cell membrane through fusion-loop insertion)  
 IT Secondary structure  
 (protein; conformational changes of **dengue** virus envelope protein E enhances its fusion to cell membrane through fusion-loop insertion)

REFERENCE COUNT: 51 THERE ARE 51 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 4 OF 12 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 6

ACCESSION NUMBER: 2003:819376 CAPLUS

DOCUMENT NUMBER: 140:90406

TITLE: **Dengue** virus M protein contains a proapoptotic sequence referred to as ApoptoM  
 AUTHOR(S): Catteau, Adeline; Kalinina, Olga; Wagner, Marie-Christine; Deubel, Vincent; Courageot, Marie-Pierre; Despres, Philippe

CORPORATE SOURCE: Unite Postulante des Interactions Moleculaires  
 Flavivirus-Hotes, Institut Pasteur, Paris, 75724/15, Fr.

SOURCE: Journal of General Virology (2003), 84(10), 2781-2793  
 CODEN: JGVIAY; ISSN: 0022-1317

PUBLISHER: Society for General Microbiology

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 19 Oct 2003

AB The induction of apoptotic cell death is a prominent cytopathic effect of dengue (DEN) viruses. One of the key questions to be addressed is which viral components induce apoptosis in DEN virus-infected cells. This study

investigated whether the small membrane (M) protein was involved in the induction of apoptosis by DEN virus. This was addressed by using a series of enhanced green fluorescent protein-fused DEN proteins. Evidence is provided that intracellular production of the M ectodomains (residues M-1 to M-40) of all four DEN serotypes triggered apoptosis in host cells such as mouse neuroblastoma Neuro 2a and human hepatoma HepG2 cells. The M ectodomains of the wild-type strains of Japanese encephalitis, West Nile and yellow fever viruses also had proapoptotic properties. The export of the M ectodomain from the Golgi apparatus to the plasma membrane appeared to be essential for the initiation of apoptosis. The study found that anti-apoptosis protein Bcl-2 protected HepG2 cells against the death-promoting activity of the DEN M ectodomain. This suggests that the M ectodomain exerts its cytotoxic effects by activating a mitochondrial apoptotic pathway. The cytotoxicity of the DEN M ectodomain reflected the intrinsic proapoptotic properties of the nine carboxy-terminal amino acids (residues M-32 to M-40) designated ApoptoM. Residue M-36 was unique in that it modulated the death-promoting activity of the M ectodomain. Defining the ApoptoM-activated signalling pathways leading to apoptosis will provide the basis for studying how the M protein might play a key role in the fate of the flavivirus-infected cells.

CC 10-1 (Microbial, Algal, and Fungal Biochemistry)

ST **dengue** virus M protein apoptosis ApoptoM

IT Proteins

RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(Bcl-2; **dengue** virus M protein contains proapoptotic sequence referred to as ApoptoM)

IT Animal cell line

(Hep G2; **dengue** virus M protein contains proapoptotic sequence referred to as ApoptoM)

IT Proteins

RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(M (matrix); **dengue** virus M protein contains proapoptotic sequence referred to as ApoptoM)

IT Protein motifs

(M **ectodomain**; **dengue** virus M protein contains proapoptotic sequence referred to as ApoptoM)

IT Animal cell line

(N2A; **dengue** virus M protein contains proapoptotic sequence referred to as ApoptoM)

IT Apoptosis

Cell membrane

**Dengue** virus

Golgi apparatus

Human

Mitochondria

Protein sequences

Signal transduction, biological

(**dengue** virus M protein contains proapoptotic sequence referred to as ApoptoM)

REFERENCE COUNT: 47 THERE ARE 47 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 5 OF 12 CAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2005:325740 CAPLUS

DOCUMENT NUMBER: 142:385982

TITLE: Small peptides having apoptotic activities and their applications

INVENTOR(S): Despres, Philippe; Catteau, Adeline

PATENT ASSIGNEE(S): Institut Pasteur, Fr.

SOURCE: U.S. Pat. Appl. Publ., 43 pp., Cont.-in-part of U.S.

Ser. No. 311,213.

CODEN: USXXCO

DOCUMENT TYPE:

Patent

LANGUAGE:

English

FAMILY ACC. NUM. COUNT: 2

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2005080231	A1	20050414	US 2003-608147	20030630
WO 2001096376	A2	20011220	WO 2001-IB1570	20010618
WO 2001096376	A3	20030313		
WO 2001096376	C2	20031023		
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 2004101862	A1	20040527	US 2003-311213	20030519
US 2004049016	A1	20040311	US 2003-634895	20030806
PRIORITY APPLN. INFO.:			US 2000-212129P	P 20000616
			WO 2001-IB1570	W 20010618
			US 2003-311213	A2 20030519
			US 2001-881710	A3 20010618

OTHER SOURCE(S): MARPAT 142:385982

ED Entered STN: 15 Apr 2005

AB The present invention relates to nine residue peptides (M32-40) from flavivirus M ectodomain able to modulate specifically the apoptotic activity of diverse flavivirus, to pharmaceutical composition comprising the same and their use for the treatment and/or the prevention of flavivirus-linked infections and cancers. Dengue virus M ectodomain peptide fusion protein containing the M precursor translocation signal sequence induced apoptosis in mouse neuroblastoma Neuro 2a and human hepatoma HepG2 cancer cells.

IC ICM C07K007-08

ICS C07K007-06

INCL 530329000; 530330000

CC 1-6 (Pharmacology)

Section cross-reference(s): 10

ST apoptotic peptide flavivirus M **ectodomain** cancer treatment; infection flavivirus treatment apoptotic peptide; **dengue virus M ectodomain** peptide apoptosis neuroblastoma hepatoma

IT Plasmids

((95-114)EGFP(M32-M40)DEN-2; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)

IT Proteins

RL: BSU (Biological study, unclassified); BIOL (Biological study) (Bcl-2, blocking apoptotic effect of M **ectodomain** peptides; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)

IT CD antigens

RL: BSU (Biological study, unclassified); BIOL (Biological study) (CD72, fragment as membrane-anchoring signal peptide targeting glycoproteins to plasma membrane; small apoptotic peptides from

flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)

IT Proteins

RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(M (matrix), apoptotic peptides of, of flavivirus; small apoptotic  
peptides from flavivirus M **ectodomain** for treating and  
preventing flavivirus-linked infections and cancers)

IT Plasmids

(Trip.quadrature.U3CMV(95-114)EGFP(237-245)DEN-2; small apoptotic  
peptides from flavivirus M **ectodomain** for treating and  
preventing flavivirus-linked infections and cancers)

IT Structure-activity relationship

(apoptosis-inducing, of **dengue** virus M protein fragments;  
small apoptotic peptides from flavivirus M **ectodomain** for  
treating and preventing flavivirus-linked infections and cancers)

IT Peptides, biological studies

RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified);  
PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological  
study); PREP (Preparation); USES (Uses)  
(apoptotic; small apoptotic peptides from flavivirus M  
**ectodomain** for treating and preventing flavivirus-linked  
infections and cancers)

IT Ligands

RL: BSU (Biological study, unclassified); PAC (Pharmacological activity);  
THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
(binding to target cancer cells, pharmaceutical composition comprising  
apoptotic peptide and; small apoptotic peptides from flavivirus M  
**ectodomain** for treating and preventing flavivirus-linked  
infections and cancers)

IT Samples

(biol., anal. of, for flavivirus infection detection; small apoptotic  
peptides from flavivirus M **ectodomain** for treating and  
preventing flavivirus-linked infections and cancers)

IT Proteins

RL: BSU (Biological study, unclassified); BUU (Biological use,  
unclassified); PAC (Pharmacological activity); THU (Therapeutic use); BIOL  
(Biological study); USES (Uses)  
(conjugates, with apoptotic peptide; small apoptotic peptides from  
flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)

IT Peptides, biological studies

RL: BSU (Biological study, unclassified); BUU (Biological use,  
unclassified); PAC (Pharmacological activity); THU (Therapeutic use); BIOL  
(Biological study); USES (Uses)  
(conjugates, with carrier protein or nonpeptide or support; small  
apoptotic peptides from flavivirus M **ectodomain** for treating  
and preventing flavivirus-linked infections and cancers)

IT Culture media

(detection of flavivirus antigens in; small apoptotic peptides from  
flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)

IT Polynucleotides

RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified);  
PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological  
study); PREP (Preparation); USES (Uses)  
(encoding apoptotic peptides; small apoptotic peptides from flavivirus  
M **ectodomain** for treating and preventing flavivirus-linked  
infections and cancers)

IT Blood analysis

(flavivirus infection detection in; small apoptotic peptides from

flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)

- IT Gene  
RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified);  
PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological  
study); PREP (Preparation); USES (Uses)  
(for apoptotic peptides; small apoptotic peptides from flavivirus M  
**ectodomain** for treating and preventing flavivirus-linked  
infections and cancers)
- IT Proteins  
RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified);  
BIOL (Biological study); PREP (Preparation)  
(green fluorescent, enhanced, marker gene for, recombinant vector  
containing; small apoptotic peptides from flavivirus M **ectodomain**  
for treating and preventing flavivirus-linked infections and cancers)
- IT Carcinoma  
(hepatocellular, induction of apoptosis in; small apoptotic peptides  
from flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)
- IT Liver, neoplasm  
(hepatoma, induction of apoptosis in; small apoptotic peptides from  
flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)
- IT Cell  
(host, transformed with recombinant vector; small apoptotic peptides  
from flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)
- IT Antibodies and Immunoglobulins  
RL: ARG (Analytical reagent use); BSU (Biological study, unclassified);  
DGN (Diagnostic use); ANST (Analytical study); BIOL (Biological study);  
USES (Uses)  
(labeled; small apoptotic peptides from flavivirus M **ectodomain**  
for treating and preventing flavivirus-linked infections and cancers)
- IT Antibodies and Immunoglobulins  
RL: ARG (Analytical reagent use); BPN (Biosynthetic preparation); BSU  
(Biological study, unclassified); DGN (Diagnostic use); ANST (Analytical  
study); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(monoclonal, to apoptotic peptide; small apoptotic peptides from  
flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)
- IT Nerve, neoplasm  
(neuroblastoma, induction of apoptosis in; small apoptotic peptides  
from flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)
- IT Diagnosis  
(of flavivirus infection, direct detection method for; small apoptotic  
peptides from flavivirus M **ectodomain** for treating and  
preventing flavivirus-linked infections and cancers)
- IT Antigens  
RL: ANT (Analyte); BSU (Biological study, unclassified); DGN (Diagnostic  
use); ANST (Analytical study); BIOL (Biological study); USES (Uses)  
(of flavivirus, detection in culture medium or biol. sample; small  
apoptotic peptides from flavivirus M **ectodomain** for treating  
and preventing flavivirus-linked infections and cancers)
- IT Mitochondrial membrane potential  
(peptide fusion protein disruption of; small apoptotic peptides from  
flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)
- IT Reactive oxygen species  
RL: MSC (Miscellaneous)

- (peptide fusion protein triggering of apoptotic pathway not involving;  
small apoptotic peptides from flavivirus M **ectodomain** for  
treating and preventing flavivirus-linked infections and cancers)
- IT Proteins  
RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(prM (premembrane), translocation signal peptide of, recombinant vector  
further comprising; small apoptotic peptides from flavivirus M  
**ectodomain** for treating and preventing flavivirus-linked  
infections and cancers)
- IT Cell membrane  
(recombinant vector further comprising membrane-anchoring signal  
peptide targeting glycoproteins to; small apoptotic peptides from  
flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)
- IT Glycoproteins  
RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(recombinant vector further comprising membrane-anchoring signal  
peptide targeting, to plasma membrane; small apoptotic peptides from  
flavivirus M **ectodomain** for treating and preventing  
flavivirus-linked infections and cancers)
- IT Proteins  
RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified);  
PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological  
study); PREP (Preparation); USES (Uses)  
(secretory pathway-targeting, recombinant vector further comprising;  
small apoptotic peptides from flavivirus M **ectodomain** for  
treating and preventing flavivirus-linked infections and cancers)
- IT Endoplasmic reticulum  
(signal peptide targeting, recombinant vector further comprising; small  
apoptotic peptides from flavivirus M **ectodomain** for treating  
and preventing flavivirus-linked infections and cancers)
- IT Antitumor agents  
Apoptosis  
Bioassay  
Dengue virus  
Dengue virus 1  
Dengue virus 2  
Dengue virus 3  
Dengue virus 4  
Drug delivery systems  
Drug screening  
Flavivirus  
Gene therapy  
Genetic markers  
Genetic vectors  
Human  
Immunoassay  
Japanese encephalitis virus  
Neoplasm  
Prophylaxis  
West Nile virus  
Yellow fever virus  
(small apoptotic peptides from flavivirus M **ectodomain** for  
treating and preventing flavivirus-linked infections and cancers)
- IT Fusion proteins (chimeric proteins)  
RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified);  
PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological  
study); PREP (Preparation); USES (Uses)  
(small apoptotic peptides from flavivirus M **ectodomain** for  
treating and preventing flavivirus-linked infections and cancers)



- IT Antibodies and Immunoglobulins  
 RL: ANT (Analyte); ARG (Analytical reagent use); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); DGN (Diagnostic use); ANST (Analytical study); BIOL (Biological study); PREP (Preparation); USES (Uses)  
 (to apoptotic peptide; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT Signal peptides  
 RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
 (translocation, recombinant vector further comprising; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT Infection  
 (viral, flavivirus, prevention and treatment of; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849707-92-2  
 RL: BSU (Biological study, unclassified); PAC (Pharmacological activity); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (M peptide of Japanese encephalitis virus, proapoptotic activity of; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849707-91-1  
 RL: BSU (Biological study, unclassified); PAC (Pharmacological activity); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (M peptide of West Nile virus, proapoptotic activity of; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849707-87-5  
 RL: BSU (Biological study, unclassified); PAC (Pharmacological activity); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (M peptide of **dengue** virus 1, proapoptotic activity of; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849707-88-6  
 RL: BSU (Biological study, unclassified); PAC (Pharmacological activity); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (M peptide of **dengue** virus 3, proapoptotic activity of; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849707-89-7  
 RL: BSU (Biological study, unclassified); PAC (Pharmacological activity); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (M peptide of **dengue** virus 4, proapoptotic activity of; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849707-90-0  
 RL: BSU (Biological study, unclassified); PAC (Pharmacological activity); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (M peptide of yellow fever virus 17D, proapoptotic activity of; small apoptotic peptides from flavivirus M **ectodomain** for treating

- and preventing flavivirus-linked infections and cancers)
- IT 849707-84-2 849707-85-3 849707-86-4  
 RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
 (M protein fragment, apoptosis induction in relation to; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849612-62-0 849612-63-1 849612-64-2 849612-65-3 849707-93-3  
 849707-94-4  
 RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
 (dengue virus M protein fragment, apoptosis induction in relation to; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 169592-56-7, Caspase-3 180189-96-2, Caspase-9 186322-81-6, Caspase  
 RL: BSU (Biological study, unclassified); BIOL (Biological study)  
 (inhibitor protecting against proapoptotic effects of M **ectodomain**; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 7782-44-7D, Oxygen, reactive species  
 RL: MSC (Miscellaneous)  
 (peptide fusion protein triggering of apoptotic pathway not involving; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849612-60-8D, fusion proteins with enhanced green fluorescent protein and M **ectodomain** peptide  
 RL: BSU (Biological study, unclassified); PAC (Pharmacological activity); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (prM translocation signal sequence, proapoptotic activity of and caspase activation by; small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849612-61-9  
 RL: BSU (Biological study, unclassified); PAC (Pharmacological activity); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)
- IT 849645-65-4  
 RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
 (small apoptotic peptides from flavivirus M **ectodomain** for treating and preventing flavivirus-linked infections and cancers)

L9 ANSWER 6 OF 12 MEDLINE on STN DUPLICATE 4  
 ACCESSION NUMBER: 2003532469 MEDLINE  
 DOCUMENT NUMBER: PubMed ID: 14610213  
 TITLE: Replication-defective adenoviral vaccine vector for the induction of immune responses to **dengue** virus type 2.  
 AUTHOR: Jaiswal Smita; Khanna Navin; Swaminathan S  
 CORPORATE SOURCE: International Centre for Genetic Engineering and Biotechnology, New Delhi 110067, India.  
 SOURCE: Journal of virology, (2003 Dec) Vol. 77, No. 23, pp.

12907-13.

Journal code: 0113724. ISSN: 0022-538X.

PUB. COUNTRY: United States  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
ENTRY MONTH: 200401  
ENTRY DATE: Entered STN: 13 Nov 2003  
Last Updated on STN: 6 Jan 2004  
Entered Medline: 5 Jan 2004

AB A recombinant replication-defective adenovirus vector that can overexpress the **ectodomain** of the envelope protein of **dengue** virus type 2 (NGC strain) has been constructed. This virus was immunogenic in mice and elicited **dengue** virus type 2 specific B- and T-cell responses. Sera from immunized mice contained neutralizing antibodies that could specifically recognize **dengue** virus type 2 and neutralize its infectivity in vitro, indicating that this approach has the potential to confer protective immunity. In vitro stimulation of splenocytes (from immunized mice) with **dengue** virus type 2 resulted in a significant proliferative response accompanied by the production of high levels of gamma interferon but did not show significant changes in interleukin-4 levels. This is suggestive of a Th1-like response (considered to be important in the maturation of cytotoxic T lymphocytes that are essential for the elimination of virus-infected cells). The data show that adenovirus vectors offer a promising alternative strategy for the development of **dengue** virus vaccines.

CT Adenoviridae: GE, genetics  
\*Adenoviridae: IM, immunology  
Adenoviridae: PH, physiology  
Animals  
Antibodies, Viral: BI, biosynthesis  
Cell Line  
Cricetinae  
Defective Viruses: GE, genetics  
\*Defective Viruses: IM, immunology  
Defective Viruses: PH, physiology  
\*Dengue Virus: IM, immunology  
Electrophoresis, Polyacrylamide Gel  
\*Genetic Vectors  
Humans  
Interferon Type II: BI, biosynthesis  
Interleukin-4: BI, biosynthesis  
Mice  
Neutralization Tests  
Research Support, Non-U.S. Gov't  
Virus Replication

L9 ANSWER 7 OF 12 MEDLINE on STN DUPLICATE 5  
ACCESSION NUMBER: 2003273060 MEDLINE  
DOCUMENT NUMBER: PubMed ID: 12759475  
TITLE: A ligand-binding pocket in the **dengue** virus envelope glycoprotein.  
AUTHOR: Modis Yorgo; Ogata Steven; Clements David; Harrison Stephen C  
CORPORATE SOURCE: Howard Hughes Medical Institute, Children's Hospital and Harvard Medical School, 320 Longwood Avenue, Boston, MA 02115, USA.  
CONTRACT NUMBER: CA13202 (NCI)  
SOURCE: Proceedings of the National Academy of Sciences of the

United States of America, (2003 Jun 10) Vol. 100, No. 12,  
pp. 6986-91. Electronic Publication: 2003-05-20.  
Journal code: 7505876. ISSN: 0027-8424.

PUB. COUNTRY: United States  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
OTHER SOURCE: PDB-1OAM; PDB-1OAN  
ENTRY MONTH: 200307  
ENTRY DATE: Entered STN: 12 Jun 2003  
Last Updated on STN: 23 Jul 2003  
Entered Medline: 22 Jul 2003

AB **Dengue** virus is an emerging global health threat. Its major envelope glycoprotein, E, mediates viral attachment and entry by membrane fusion. A crystal structure of the soluble **ectodomain** of E from **dengue** virus type 2 reveals a hydrophobic pocket lined by residues that influence the pH threshold for fusion. The pocket, which accepts a hydrophobic ligand, opens and closes through a conformational shift in a beta-hairpin at the interface between two domains. These features point to a structural pathway for the fusion-activating transition and suggest a strategy for finding small-molecule inhibitors of **dengue** and other flaviviruses.

CT Binding Sites  
Crystallography, X-Ray  
Dengue Virus: GE, genetics  
\*Dengue Virus: ME, metabolism  
Dengue Virus: PY, pathogenicity  
Dengue Virus: PH, physiology  
Dimerization  
Humans  
Ligands  
Membrane Fusion  
Models, Molecular  
Peptide Fragments: CH, chemistry  
Peptide Fragments: GE, genetics  
Peptide Fragments: ME, metabolism  
Protein Structure, Quaternary  
Protein Subunits  
Recombinant Proteins: CH, chemistry  
Recombinant Proteins: GE, genetics  
Recombinant Proteins: ME, metabolism  
Research Support, Non-U.S. Gov't  
Research Support, U.S. Gov't, P.H.S.  
\*Viral Envelope Proteins: CH, chemistry  
Viral Envelope Proteins: GE, genetics  
\*Viral Envelope Proteins: ME, metabolism  
Virus Assembly

L9 ANSWER 8 OF 12 MEDLINE on STN DUPLICATE 7  
ACCESSION NUMBER: 2003256723 MEDLINE  
DOCUMENT NUMBER: PubMed ID: 12783086  
TITLE: Dendritic-cell-specific ICAM3-grabbing non-integrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived **dengue** viruses.  
AUTHOR: Navarro-Sanchez Erika; Altmeyer Ralf; Amara Ali; Schwartz Olivier; Fieschi Franck; Virelizier Jean-Louis; Arenzana-Seisdedos Fernando; Despres Philippe  
CORPORATE SOURCE: Interactions Moleculaires Flavivirus-Hotes, 25 Rue du Dr Roux, 75724 Paris, France.  
SOURCE: EMBO reports, (2003 Jul) Vol. 4, No. 7, pp. 723-8.

Journal code: 100963049. ISSN: 1469-221X.

PUB. COUNTRY: England: United Kingdom  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
ENTRY MONTH: 200404  
ENTRY DATE: Entered STN: 4 Jun 2003  
Last Updated on STN: 21 Apr 2004  
Entered Medline: 20 Apr 2004

AB Dengue virus (DV) is a mosquito-borne flavivirus that causes haemorrhagic fever in humans. DV primarily targets immature dendritic cells (DCs) after a bite by an infected mosquito vector. Here, we analysed the interactions between DV and human-monocyte-derived DCs at the level of virus entry. We show that the DC-specific ICAM3-grabbing non-integrin (DC-SIGN) molecule, a cell-surface, mannose-specific, C-type lectin, binds mosquito-cell-derived DVs and allows viral replication. Conclusive evidence for the involvement of DC-SIGN in DV infection was obtained by the inhibition of viral infection by anti-DC-SIGN antibodies and by the soluble tetrameric ectodomain of DC-SIGN. Our data show that DC-SIGN functions as a DV-binding lectin by interacting with the DV envelope glycoprotein. Mosquito-cell-derived DVs may have differential infectivity for DC-SIGN-expressing cells. We suggest that the differential use of DC-SIGN by viral envelope glycoproteins may account for the immunopathogenesis of DVs.

CT Animals  
Antibodies, Monoclonal: IM, immunology  
Antibodies, Monoclonal: PD, pharmacology  
\*Antigens, CD: ME, metabolism  
Cell Adhesion Molecules: AI, antagonists & inhibitors  
Cell Adhesion Molecules: IM, immunology  
\*Cell Adhesion Molecules: ME, metabolism  
Cells, Cultured  
\*Culicidae: VI, virology  
Dendritic Cells: CY, cytology  
\*Dendritic Cells: ME, metabolism  
\*Dendritic Cells: VI, virology  
\*Dengue Virus: PH, physiology  
Fluorescent Antibody Technique, Direct  
Humans  
Hydrogen-Ion Concentration  
Lectins, C-Type: AI, antagonists & inhibitors  
Lectins, C-Type: IM, immunology  
\*Lectins, C-Type: ME, metabolism  
Monocytes: CY, cytology  
Monocytes: ME, metabolism  
Monocytes: VI, virology  
Receptors, Cell Surface: AI, antagonists & inhibitors  
Receptors, Cell Surface: IM, immunology  
\*Receptors, Cell Surface: ME, metabolism  
Research Support, Non-U.S. Gov't

L9 ANSWER 9 OF 12 MEDLINE on STN DUPLICATE 8  
ACCESSION NUMBER: 1999139041 MEDLINE  
DOCUMENT NUMBER: PubMed ID: 9971841  
TITLE: PrM- and cell-binding domains of the dengue virus E protein.  
AUTHOR: Wang S; He R; Anderson R  
CORPORATE SOURCE: Department of Microbiology and Immunology, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada.  
SOURCE: Journal of virology, (1999 Mar) Vol. 73, No. 3, pp.

2547-51.

Journal code: 0113724. ISSN: 0022-538X.

PUB. COUNTRY: United States  
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
LANGUAGE: English  
FILE SEGMENT: Priority Journals  
ENTRY MONTH: 199903  
ENTRY DATE: Entered STN: 16 Mar 1999  
Last Updated on STN: 16 Mar 1999  
Entered Medline: 4 Mar 1999

AB The E-prM proteins of flaviviruses are unusual complexes which play important roles in virus assembly and fusion modulation and in potential immunity-inducing vaccines. Despite their importance, little is known about the biogenesis and structural organization of E-prM complexes. Pulse-chase radiolabeling of **dengue** virus-infected Vero cells demonstrated a rapid interassociation of E and prM proteins, and sucrose gradient sedimentation analysis suggested that E-prM complexes progressed from simple heteromers to more densely sedimenting structures indicating increased multimerization. E-prM heteromers of even higher complexity were observed in virus particles, suggesting an intracellular assembly process which results in the networking of E-prM subunits into a lattice-like structure found in virus particles. Trypsin cleavage of E-prM-containing virus particles resulted in the release of a soluble 45-kDa fragment of the E protein which retained cell-binding activity. The results suggest that E-prM interactions in **dengue** virus particles are largely mediated by domains in the carboxy-terminal anchoring domain of E, while cell-binding activity is retained in a trypsin-releasable **ectodomain** of the E protein.

CT Animals  
Binding Sites  
Cercopithecus aethiops  
\*Dengue Virus: PH, physiology  
\*Receptors, Virus: PH, physiology  
Research Support, Non-U.S. Gov't  
Trypsin: PD, pharmacology  
Vero Cells  
\*Viral Proteins: PH, physiology  
Virion: PH, physiology

L9 ANSWER 10 OF 12 WPIX COPYRIGHT 2006 THE THOMSON CORP on STN  
ACCESSION NUMBER: 2006-332034 [34] WPIX  
DOC. NO. CPI: C2006-109493  
TITLE: Attenuated Flavivirus, useful for preparing a vaccine against Japanese Encephalitis Virus, West Nile Virus or other disease-causing Flavivirus, comprises a membrane protein mutation.  
DERWENT CLASS: B04 D16  
INVENTOR(S): CATALAN, J A; GUIRAKHOO, F; LIU, J; MONATH, T P; PUGACHEV, K V  
PATENT ASSIGNEE(S): (ACAM-N) ACAMBIS INC  
COUNTRY COUNT: 112  
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2006044857	A2	20060427	(200634)*	EN	122
RW: AT BE BG BW CH CY CZ DE DK EA EE ES FI FR GB GH GM GR HU IE IS IT					
KE LS LT LU LV MC MW MZ NA NL OA PL PT RO SD SE SI SK SL SZ TR TZ					
UG ZM ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE					

DK DM DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG  
 KM KP KR KZ LC LK LR LS LT LU LV LY MA MD MG MK MN MW MX MZ NA NG  
 NI NO NZ OM PG PH PL PT RO RU SC SD SE SG SK SL SM SY TJ TM TN TR  
 TT TZ UA UG US UZ VC VN YU ZA ZM ZW

## APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2006044857	A2	WO 2005-US37369	20051019

PRIORITY APPLN. INFO: US 2005-718923P 20050919; US  
 2004-620466P 20041020; US  
 2004-620948P 20041021; US  
 2005-674415P 20050424; US  
 2005-674546P 20050425

AB WO2006044857 A UPAB: 20060526

NOVELTY - A recombinant Flavivirus comprising a membrane protein mutation, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for:

- (1) a vaccine composition comprising the Flavivirus above and a pharmaceutical carrier or diluent;
- (2) inducing an immune response to a Flavivirus in a patient by administering the vaccine composition;
- (3) producing a vaccine composition comprising a recombinant Flavivirus by introducing a mutation into the membrane protein of the Flavivirus;
- (4) a nucleic acid molecule corresponding to the genome of the Flavivirus comprising a membrane protein mutation; and
- (5) manufacturing the Flavivirus with a membrane protein mutation by introducing a nucleic acid molecule corresponding to the virus genome into cells and isolating Flavivirus produced in the cells from the cells or the culture supernatant.

ACTIVITY - Virucide.

No biological data given.

MECHANISM OF ACTION - Vaccine.

USE - The Flavivirus, vaccine composition and method are useful for inducing an immune response to a Flavivirus. The virus is preferably a Japanese encephalitis virus, West Nile virus, yellow fever virus, dengue virus, St. Louis encephalitis virus, Murray valley encephalitis virus, or tick-borne encephalitis virus. (All claimed).  
 Dwg.0/9

L9 ANSWER 11 OF 12 WPIX COPYRIGHT 2006 THE THOMSON CORP on STN

ACCESSION NUMBER: 2004-097654 [10] WPIX

DOC. NO. NON-CPI: N2004-077771

DOC. NO. CPI: C2004-040532

TITLE: Identifying antibody that inhibits pathogenicity of infectious microorganism such as an orthopox virus using a programmed computer.

DERWENT CLASS: B04 D16 S05 T01

INVENTOR(S): RECHE, P; REINHERZ, E L

PATENT ASSIGNEE(S): (RECH-I) RECHE P; (REIN-I) REINHERZ E L; (DAND) DANA FARBER CANCER INST INC

COUNTRY COUNT: 102

PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
-----------	------	------	------	----	----

US 2003229454 A1 20031211 (200410)\* 25  
 WO 2004040398 A2 20040513 (200439) EN  
 RW: AT BE BG CH CY CZ DE DK EA EE ES FI FR GB GH GM GR HU IE IT KE LS  
 LU MC MW MZ NL OA PT RO SD SE SI SK SL SZ TR TZ UG ZM ZW  
 W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK  
 DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR  
 KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NI NO NZ OM PH PL  
 PT RO RU SC SD SE SG SK SL TJ TM TN TR TT TZ UA UG US UZ VC VN ZA  
 ZM ZW  
 AU 2003299469 A1 20040525 (200468)

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2003229454	A1 Provisional	US 2002-380055P	20020506
	Provisional	US 2003-453649P	20030311
		US 2003-429685	20030505
WO 2004040398	A2	WO 2003-US13982	20030505
AU 2003299469	A1	AU 2003-299469	20030505

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2003299469	A1 Based on	WO 2004040398

PRIORITY APPLN. INFO: US 2003-429685 20030505; US  
 2002-380055P 20020506; US  
 2003-453649P 20030311

AB US2003229454 A UPAB: 20040210  
 NOVELTY - Identifying antibody using programmed computer comprising inputting into input device a nucleic acid sequence of open reading frames (ORF) in genome of infectious microorganism, screening nucleic acid sequence to identify ORF encoding protein expressed on surface of microorganism, producing antibody that binds to **ectodomain** of protein and determining whether antibody inhibits pathogenicity of infectious microorganism, is new.

DETAILED DESCRIPTION - A computer-based method (M1) of identifying an antibody that inhibits infection, which uses a programmed computer comprising a processor and an input device, involves providing the nucleic acid sequence of several open reading frames (ORF) in the genome of an infectious microorganism, inputting to the input device the nucleic acid sequence, screening the nucleic acid sequence using the processor, to identify an open reading frame encoding a protein that is predicted to be expressed on the surface of the infectious microorganism, producing an antibody that binds to the **ectodomain** of the protein, and determining whether the antibody inhibits the pathogenicity of the infectious microorganism.

INDEPENDENT CLAIMS are also included for:

(1) manufacturing (M2) a compound, involves performing (M1) and after determining that the antibody inhibits the pathogenicity of the infectious microorganism, and manufacturing a compound comprising at least a portion of the **ectodomain**;

(2) manufacturing (M3) an antibody, involves performing (M1) and after determining that the antibody inhibits the pathogenicity of the infectious microorganism, manufacturing the antibody;

(3) a compound manufactured by (M2);

(4) inducing an immune response in an animal, involves performing (M1) and after determining that the antibody inhibits pathogenicity of the



infectious microorganism, administering a compound comprising at least a portion of the **ectodomain** to an animal susceptible to infection with the infectious microorganism;

(5) treating (M4), involves performing (M1) and after determining that the antibody inhibits, pathogenicity of the infectious microorganism, administering the antibody to an animal;

(6) an antibody manufactured by (M3);

(7) a monoclonal antibody (I) that binds to a protein encoded by the genome of variola virus or a vaccinia virus, where the protein is a protein that is expressed on the surface of the virus or on the surface of a cell infected with the virus; and

(8) a humanized antibody derived from (I).

ACTIVITY - Antimicrobial.

A group of C57BL/6 (B6) mice (n=5) were injected intraperitoneally with a cocktail containing 200 micro g of each of the 11D7, 13E8, and 7D11 monoclonal antibodies. A second group of B6 mice (n=5) were injected intraperitoneally with 600 micro g of a control monoclonal antibody (1A3). Six hours later, the mice in both groups were challenged intranasally with an LD(50) dose (104 PFU) of vaccinia virus (strain WR). On the ninth day after infection, all mice were sacrificed and the amount of virus in their lungs (in PFU/ml) was measured. The animals injected with the cocktail of antibodies showed a dramatic decrease in viral titer compared with the control mice. No virus was detectable in the lungs of animals 6, 7, 9 and 10. About 102 PFU/ml were detected in the lungs of the animals because 102 PFU/ml was the detection limit of the assay. While all the control animals showed severe morbidity starting on day 7 and loss of fat pads, all the experimental animals appeared clinically normal, retained their fat pads, and had no lung pathology at the time of sacrifice.

MECHANISM OF ACTION - Inhibits pathogenicity of microorganisms (claimed).

USE - (M1) is useful for identifying an antibody that inhibits pathogenicity of a infectious microorganism. The infectious microorganism is a virus (such as variola major/minor virus or vaccinia virus). The virus is chosen from hepatitis virus A-E, human papilloma virus, human immunodeficiency virus 1, human T cell lymphotropic virus 1, Herpes virus, Dengue virus 1-4, Ebola virus, Marburg virus, Lassa virus, Machupo virus and influenza virus. The infectious microorganism is a bacterium. The bacterium is Mycobacterium tuberculosis, Mycobacterium leprae, Salmonella bacterium (such as Salmonella typhimurium or Salmonella typhi) or Yersinia pestis. The bacterium is also chosen from Bacillus anthracis, Clostridium botulinum, Francisella tularensis, Corynebacterium diphtheriae, Vibrio cholerae and Escherichia coli. The infectious microorganism is a protozoan parasite such as a malarial parasite or Leishmania. (M1) is useful for manufacturing a compound comprising at least a portion of the **ectodomain**. (M1) is useful for manufacturing an antibody which inhibits pathogenicity of an infectious microorganism. (M1) is also useful for inducing an immune response in an animal. The immune response is a protective immune response. (M4) is useful for treating by administering an antibody that inhibits pathogenicity of the infectious microorganism. The infectious microorganism is a virus. The virus is an orthopox virus (such as variola or vaccinia virus). The protein is a smallpox growth factor (SPGF) or a VGF (vaccinia growth factor). The monoclonal antibody is the 3D4R-13E8 monoclonal antibody (ATCC Accession Number PTA-5040) or 3D4R-11D7 monoclonal antibody (ATCC Accession Number PTA-5039). The method further involves administering to the animal one or more additional antibodies, where one or more additional antibodies bind to a protein encoded by the infectious microorganism such as an orthopox virus (all claimed).

Dwg.5/6

L9 ANSWER 12 OF 12 WPIX COPYRIGHT 2006 THE THOMSON CORP on STN  
 ACCESSION NUMBER: 2002-139706 [18] WPIX  
 DOC. NO. CPI: C2002-043009  
 TITLE: Novel apoptosis inducing polypeptide fragments of  
**Dengue** virus-1 or 2 M protein, useful for  
 inducing apoptosis in a cell of a human patient suffering  
 from cancer or flavivirus infection.  
 DERWENT CLASS: B04 D16  
 INVENTOR(S): CATTEAU, A; COURAGEOT, M; DESPRES, P; DEUBEL, V  
 PATENT ASSIGNEE(S): (INSP) INST PASTEUR; (CATT-I) CATTEAU A; (COUR-I)  
 COURAGEOT M; (DESP-I) DESPRES P; (DEUB-I) DEUBEL V  
 COUNTRY COUNT: 95  
 PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2001096376	A2	20011220	(200218)*	EN	45
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZW W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW					
AU 2001082388	A	20011224	(200227)		
US 2002086403	A1	20020704	(200247)		
EP 1311539	A2	20030521	(200334)	EN	
R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI TR					
US 6673895	B2	20040106	(200411)		
JP 2004503233	W	20040205	(200412)		78
US 2004049016	A1	20040311	(200419)		
US 2004101862	A1	20040527	(200435)		
US 2005080231	A1	20050414	(200526)		

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2001096376	A2	WO 2001-IB1570	20010618
AU 2001082388	A	AU 2001-82388	20010618
US 2002086403	A1 Provisional	US 2000-212129P	20000616
		US 2001-881710	20010618
EP 1311539	A2	EP 2001-961004	20010618
		WO 2001-IB1570	20010618
US 6673895	B2 Provisional	US 2000-212129P	20000616
		US 2001-881710	20010618
JP 2004503233	W	WO 2001-IB1570	20010618
		JP 2002-510516	20010618
US 2004049016	A1 Provisional Div ex	US 2000-212129P	20000616
		US 2001-881710	20010618
		US 2003-634895	20030806
US 2004101862	A1	WO 2001-IB1570	20010618
		US 2003-311213	20030519
US 2005080231	A1 Provisional	US 2000-212129P	20000616
	CIP of	WO 2001-IB1570	20010618
	CIP of	US 2003-311213	20030519
		US 2003-608147	20030630

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2001082388	A Based on	WO 2001096376
EP 1311539	A2 Based on	WO 2001096376
JP 2004503233	W Based on	WO 2001096376
US 2004049016	A1 Div ex	US 6673895

PRIORITY APPLN. INFO: US 2000-212129P 20000616; US  
 2001-881710 20010618; US  
 2003-634895 20030806; US  
 2003-311213 20030519; US  
 2003-608147 20030630

AB WO 200196376 A UPAB: 20020319

NOVELTY - An isolated polypeptide (I) having a fully defined (a) **Dengue virus (DEN)-1 M** (a membrane protein anchored in envelope surrounding the nucleocapsid of the virus) **ectodomain** sequence, (b) Den-1-C amino acid sequence of 95-114 as given in specification, or (c) DEN-2 M **ectodomain** sequence, is new.

DETAILED DESCRIPTION - (I) has a fully defined sequence of SVALAPHVGLGLETRTETWMSSEGAWKQIQKVETWALRHP (DEN-1 M **ectodomain**) (S1) or SVALVPHVGMGLETRTETWMSSEGAWKHAQRIETWILRHP (DEN-2 M **ectodomain**) (S3), or a fully defined Den-1 C amino acid sequence (S2) of 95-114 as given in the specification.

INDEPENDENT CLAIMS are also included for the following:

- (1) an isolated polynucleotide (II) which encodes (I);
- (2) a vector (III) comprising (II);
- (3) a prokaryotic or eukaryotic cell comprising (II);
- (4) a composition comprising (I) or (II) and a carrier;
- (5) an isolated polypeptide (IV) of the sequence of (S1), operably linked to (S2);
- (6) an isolated polynucleotide (V) which encodes (IV);
- (7) a vector (VI) comprising (V);
- (8) a prokaryotic or eukaryotic cell comprising (V);
- (9) a composition comprising (IV) or (V) and a carrier;
- (10) screening (M1) for peptides capable of inducing apoptosis involves introducing a recombinant protein into the cell, where the recombinant protein comprises the peptide to be screened operably linked to (S2) and detecting apoptosis in a cell;
- (11) monoclonal antibodies raised against DEN-1 or DEN-2 viral M protein;
- (12) plasmid (95-114)EGFP(M10-M40)DEN-2 deposited at the CNCM under the accession number I-2684;
- (13) plasmid pTrip Delta U3(95-114)EGFP(206-245)DEN-2 deposited at the CNCM under the accession number I-2686;
- (14) plasmid pTrip Delta U3(95-114)EGFP(206-245)DEN-1 deposited at the CNCM under the accession number I-2685; and
- (15) plasmid p(95-114)EGFP(215-245)WNV deposited at the CNCM under the accession number I-2475.

ACTIVITY - Virucide; cytostatic.

MECHANISM OF ACTION - Apoptosis inducer.

To test the pro-apoptotic activity of the DEN-2 M **ectodomain** the chimeric protein (95-114)EGFP(206-245)DEN-2 was employed. The region of the DEN-2 virus strain Jamaica corresponding the M **ectodomain** (DEN-2 polyprotein 206-245) was fused to the C-terminus of the (95-114)EGFP fusion construct. The cytotoxicity of the (95-114)EGFP(206-245)DEN-2 chimeric protein was tested by transfecting cells with FuGENE 6. The expression of the chimeric protein was observed by the autofluorescence of the EGFP and apoptotic cell death was detected visually by staining with propidium iodide as described above. Intracellular expression of the (95-114)EGFP(206-245)DEN-2 chimeric

protein resulted in cell death. DEN-2 M **ectodomain** has the ability to induce rapid apoptosis in Neuro 2a, HepG2, HeLa, and VERO cells. Apoptosis was more pronounced after transfection with plasmid (95-114)EGFP(206-245)DEN-2 than after transfection with the plasmid (95-114)EGFP(206-245) containing the sequence of the DEN-1 M **ectodomain**.

USE - (I) having a sequence of (S3) is useful for inducing apoptosis in the cell of a human patient suffering from cancer or an infection with flavivirus. (I) having a sequence of (S1) or (S3) is useful for screening for molecules which inhibit apoptosis induced by the polypeptide. The method involves introducing the polypeptide into a cell; contacting the cell containing the polypeptide, with the molecule to be screened; and detecting the presence or absence of apoptosis in the cell. The polypeptide which is introduced into the cell is operably linked to the polypeptide of sequence (S2) or (S3) or to a green fluorescent protein. Optionally, the polypeptide is not linked to a green fluorescent protein. Preferably, the polypeptide is introduced into a cell by introducing a polynucleotide (which is an expression vector capable of expressing the polypeptide in a cell) which encodes the polypeptide. (IV) when administered or delivered into a cell is useful for inducing apoptosis in the cell, preferably present in a human patient suffering from cancer or infected with flavivirus. The delivery of (IV) is carried out by delivering a polynucleotide encoding the polypeptide to the cell, where the polynucleotide is in an expression vector suitable to express the polypeptide in the cell (all claimed). The polynucleotides and polypeptides encoded by the polynucleotides as described above are useful for inducing apoptosis, and for treating patients with cancer and patients infected with flavivirus.

Dwg.0/15

L19 ANSWER 1 OF 14 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 1  
 ACCESSION NUMBER: 2004:740442 CAPLUS  
 DOCUMENT NUMBER: 141:237791  
 TITLE: Protein and cDNA **sequences** of new  
**Dengue** and West Nile viruses glycoproteins,  
 and their use in vaccinal, therapeutic and diagnostic  
 applications  
 INVENTOR(S): Tangy, Frederic; **Despres, Philippe**;  
 Combredet, Chantal; Frenkiel, Marie Pascale  
 PATENT ASSIGNEE(S): Institut Pasteur, Fr.; Centre National de la Recherche  
 Scientifique C.N.R.S.  
 SOURCE: PCT Int. Appl., 64 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
-----	----	-----	-----	-----
WO 2004076619	A2	20040910	WO 2004-IB1027	20040226
WO 2004076619	A3	20050317		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH,			
	CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD,			
	GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,			
	LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI			
RW:	BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE,			

BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU,  
 MC, NL, PT, RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN,  
 GQ, GW, ML, MR, NE, SN, TD, TG

CA 2420092	AA	20040826	CA 2003-2420092	20030226
CA 2432738	AA	20040826	CA 2003-2432738	20030620
CA 2456873	AA	20040826	CA 2004-2456873	20040226
CA 2517258	AA	20040910	CA 2004-2517258	20040226
EP 1599495	A2	20051130	EP 2004-714865	20040226

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,  
 IE, SI, LT, LV, FI, RO, MK, CY, AL, TR, BG, CZ, EE, HU, SK

BR 2004007840	A	20060214	BR 2004-7840	20040226
US 2006073164	A1	20060406	US 2005-210960	20050825

PRIORITY APPLN. INFO.: CA 2003-2420092 A 20030226  
 CA 2003-2432738 A 20030620  
 WO 2004-IB1027 W 20040226

AB The present invention relates to the development of viral vectors expressing different immunogens from the West Nile Encephalitis Virus (WNV) or the Dengue virus which are able to induce protective humoral and cellular immune responses against WNV or Dengue virus infections. More specifically, the present invention relates to three (3) antigens from WNV (the secreted envelope glycoprotein (E), the heterodimer glycoproteins (pre-M-E) and the NS1 protein) and from Dengue virus (the secreted envelope glycoprotein (e), the heterodimer glycoproteins (pre- m-e) and the NS1 protein) and their use in vaccinal, therapeutic and diagnostic applications.

L19 ANSWER 2 OF 14 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 2

ACCESSION NUMBER: 2003:848534 CAPLUS

DOCUMENT NUMBER: 140:39857

TITLE: Expression of **dengue** ApoptoM  
**sequence** results in disruption of  
 mitochondrial potential and caspase activation

AUTHOR(S): **Catteau, Adeline**; Roue, Gael; Yuste, Victor  
 J.; Susin, Santos A.; **Despres, Philippe**

CORPORATE SOURCE: Unite des Interactions Moleculaires Flavivirus-Hotes,  
 Institut Pasteur, Paris, 75015, Fr.

SOURCE: Biochimie (2003), 85(8), 789-793

CODEN: BICMBE; ISSN: 0300-9084

PUBLISHER: Editions Scientifiques et Medicales Elsevier

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Apoptotic cell death was involved as a cytopathol. mechanism in response to dengue (DEN) virus infection. Little information exists about how DEN virus replication triggers apoptosis in infected cells. We reported that a 9-residue sequence of the DEN M protein referred to as ApoptoM has proapoptotic properties in transformed and tumor cells of various origins. The aim of the present study was to investigate whether ApoptoM-induced apoptosis is associated to mitochondrial dysfunction and requires caspase activation. Intracellular expression of ApoptoM provokes the disruption of the mitochondrial transmembrane potential without subsequent generation of reactive oxygen species. We showed that ApoptoM-induced apoptosis involves the activation of a caspase-like protease pathway. Caspase-3 like activity was detected in ApoptoM-expressing cells. However, there was no role for caspase-9 in ApoptoM-mediated cell death. These data suggest that a particular mitochondrion-dependent apoptotic pathway may be involved in induction of apoptosis by ApoptoM.

REFERENCE COUNT: 12 THERE ARE 12 CITED REFERENCES AVAILABLE FOR THIS  
 RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L19 ANSWER 3 OF 14 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 3

ACCESSION NUMBER: 2002:893884 CAPLUS  
 DOCUMENT NUMBER: 138:249386  
 TITLE: Genome analysis of dengue type-1 virus isolated between 1990 and 2001 in Brazil reveals a remarkable conservation of the structural proteins but amino acid differences in the non-structural proteins  
 AUTHOR(S): Duarte dos Santos, Claudia Nunes; Rocha, Carlos Fernando S.; Cordeiro, Marli; Fragoso, Stenio P.; Rey, Felix; Deubel, Vincent; **Despres, Philippe**  
 CORPORATE SOURCE: Instituto de Biologia Molecular do Parana, Rio de Janeiro, Brazil  
 SOURCE: Virus Research (2002), 90(1-2), 197-205  
 CODEN: VIREDF; ISSN: 0168-1702  
 PUBLISHER: Elsevier Science Ltd.  
 DOCUMENT TYPE: Journal  
 LANGUAGE: English  
 AB We have investigated the genetic diversity of dengue type-1 (DEN-1) virus in Brazil. The full nucleotide sequences of three DEN-1 virus isolated from DEN fever (DF) and DEN hemorrhagic fever patients in northeastern Brazil in 1997 (BR/97) and one from a DF patient in the south of Brazil in 2001 (BR/01) were compared to that of the reference strain BR/90 obtained in the city of Rio de Janeiro in 1990. Sequence anal. showed that the structural proteins were remarkably conserved between all isolates. A total of 27 amino acid changes occurred throughout the non-structural proteins. Among them, nine amino acid substitutions were specific of BR/97 and BR/01 isolates, indicating that in situ evolution of these strains had occurred. Within the BR/97 and BR/01 samples, some amino acid substitutions have been previously identified in DEN-1 virus strains sequenced so far, suggesting that recombination events might have occurred.  
 REFERENCE COUNT: 28 THERE ARE 28 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L19 ANSWER 4 OF 14 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 4  
 ACCESSION NUMBER: 2000:881428 CAPLUS  
 DOCUMENT NUMBER: 134:39168  
 TITLE: Early detection of flaviviruses using antibodies to NS1 glycoprotein  
 INVENTOR(S): Flamand, Marie; Megret, Francoise; Alcon, Sophie; Talarmin, Antoine; **Despres, Philippe**; Deubel, Vincent  
 PATENT ASSIGNEE(S): Institut Pasteur, Fr.  
 SOURCE: PCT Int. Appl., 51 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: French  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000075665	A1	20001214	WO 2000-FR1620	20000609
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ,			

CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

FR 2794864	A1	20001215	FR 1999-7290	19990609
FR 2794865	A1	20001215	FR 1999-7361	19990610
FR 2794865	B1	20030418		
EP 1190257	A1	20020327	EP 2000-951579	20000609
EP 1190257	B1	20040825		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				
BR 2000011369	A	20030812	BR 2000-11369	20000609
AT 274701	E	20040915	AT 2000-951579	20000609
AU 776844	B2	20040923	AU 2000-64474	20000609
PT 1190257	T	20041231	PT 2000-951579	20000609
ES 2226894	T3	20050401	ES 2000-951579	20000609
ZA 2001009993	A	20020823	ZA 2001-9993	20011205
US 6870032	B1	20050322	US 2002-980839	20020621
HK 1045729	A1	20050318	HK 2002-107117	20020926
US 2005186562	A1	20050825	US 2004-17048	20041221
PRIORITY APPLN. INFO.:			FR 1999-7290	A 19990609
			FR 1999-7361	A 19990610
			WO 2000-FR1620	W 20000609
			US 2002-980839	A3 20020621

AB The invention concerns a method for early detection of a flavivirus-induced infection, comprising the detection of the flavivirus non-structural glycoprotein NS1 in a biol. sample during the clin. phase of the infection, by an immunol. method using at least two identical or different antibodies, the first antibody consisting of polyclonal or monoclonal antibodies pre-selected for their high affinity for said NS1 protein hexameric in shape.

REFERENCE COUNT: 9 THERE ARE 9 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L19 ANSWER 5 OF 14 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 6

ACCESSION NUMBER: 1993:642584 CAPLUS

DOCUMENT NUMBER: 119:242584

TITLE: Differences between cell membrane fusion activities of two dengue type-1 isolates reflect  
• modifications of viral structure

AUTHOR(S): Despres, Philippe; Frenkiel, Marie Pascale; Deubel, Vincent

CORPORATE SOURCE: Unite Arbovirus Virus Fievres Hemorragiques, Inst. Pasteur, Paris, 75724, Fr.

SOURCE: Virology (1993), 196(1), 209-19  
CODEN: VIRLAX; ISSN: 0042-6822

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The genetic diversity of dengue (DEN) virus was explored using two South American DEN-1 virus strains isolated from viremic human sera. DEN-1 virus strains BR/90 and FGA/89 were selected on the basis of their membrane fusion properties in mosquito cell cultures. Infection of mosquito cell lines with BR/90 virus strain induced a cytopathic effect characterized by syncytium formation whereas no cytopathic changes were observed with FGA/89. Cell-to-cell fusion expts. indicated that the fusogenic activity of FGA/89 required a lower pH than BR/90. Immunoreactivity anal. of the DEN-1 envelope (E) protein with monoclonal antibodies revealed a minor difference between the antigenic structures of FGA/89 and BR/90 virions. FGA/89 was less neurovirulent than BR/90 for newborn mouse. To determine the genetic origin of these modifications, the amino acid sequences of the structural proteins from these virus strains

were compared. One amino acid difference was found within the carboxy-terminal domain of protein C. Five amino acid substitutions were found in the E proteins at positions 96, 180, 297, 379, and 473. Changes at positions 96, 297, and 379 map within two overlapping antigenic domains of protein E. These limited amino acid differences in the E protein could affect the biol. properties and the antigenicity of the DEN virion.

L19 ANSWER 6 OF 14 CAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2001:923833 CAPLUS  
DOCUMENT NUMBER: 136:52711  
TITLE: Pro-apoptotic fragments of the dengue virus envelope glycoproteins  
INVENTOR(S): Despres, Philippe; Courageot, Marie-Pierre; Deubel, Vincent; Catteau, Adeline  
PATENT ASSIGNEE(S): Institut Pasteur, Fr.  
SOURCE: PCT Int. Appl., 45 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 2  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001096376	A2	20011220	WO 2001-IB1570	20010618
WO 2001096376	A3	20030313		
WO 2001096376	C2	20031023		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
CA 2412257	AA	20011220	CA 2001-2412257	20010618
US 2002086403	A1	20020704	US 2001-881710	20010618
US 6673895	B2	20040106		
EP 1311539	A2	20030521	EP 2001-961004	20010618
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR			
JP 2004503233	T2	20040205	JP 2002-510516	20010618
US 2004101862	A1	20040527	US 2003-311213	20030519
US 2005080231	A1	20050414	US 2003-608147	20030630
US 2004049016	A1	20040311	US 2003-634895	20030806
PRIORITY APPLN. INFO.:			US 2000-212129P	P 20000616
			US 2001-881710	A3 20010618
			WO 2001-IB1570	W 20010618
			US 2003-311213	A2 20030519

AB The present invention relates to pro-apoptotic fragments of the Dengue virus prM and E glycoproteins, methods of screening for mols. capable of inducing apoptosis and methods of inducing apoptosis in a cell. The M, prM and E proteins of DEN-1 or DEN-2, and monoclonal antibodies can be used for diagnostic and therapeutic agents against Flavivirus infection and cancer.

L19 ANSWER 7 OF 14

MEDLINE on STN

DUPLICATE 5

ACCESSION NUMBER: 2000461004 MEDLINE



DOCUMENT NUMBER: PubMed ID: 10964773  
 TITLE: Determinants in the envelope E protein and viral RNA helicase NS3 that influence the induction of apoptosis in response to infection with **dengue** type 1 virus.  
 AUTHOR: Duarte dos Santos C N; Frenkiel M P; Courageot M P; Rocha C F; Vazeille-Falcoz M C; Wien M W; Rey F A; Deubel V; **Despres P**  
 CORPORATE SOURCE: Departamento de Bioquimica e Biologia Molecular, Laboratorio de Expressao e Regulacao Genica, Rio de Janeiro, R.J., Brazil.  
 SOURCE: Virology, (2000 Sep 1) Vol. 274, No. 2, pp. 292-308. Journal code: 0110674. ISSN: 0042-6822.  
 PUB. COUNTRY: United States  
 DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
 LANGUAGE: English  
 FILE SEGMENT: Priority Journals  
 OTHER SOURCE: GENBANK-AF226686; GENBANK-AF226687  
 ENTRY MONTH: 200009  
 ENTRY DATE: Entered STN: 5 Oct 2000  
 Last Updated on STN: 5 Oct 2000  
 Entered Medline: 25 Sep 2000

AB One mechanism by which **dengue** (DEN) virus may cause cell death is apoptosis. In this study, we investigated whether the genetic determinants responsible for acquisition by DEN type 1 (DEN-1) virus of mouse neurovirulence interfere with the induction of apoptosis. Neurovirulent variant FGA/NA d1d was generated during the adaptation of the human isolate of DEN-1 virus strain FGA/89 to grow in newborn mouse brains and mosquito cells in vitro [Despres, P. Frenkiel, M. -P. Ceccaldi, P.-E. Duarte Dos Santos, C. and Deubel, V. (1998) J. Virol., 72: 823-829]. Genetic determinants possibly responsible for mouse neurovirulence were studied by sequencing the entire genomes of both DEN-1 viruses. Three amino acid differences in the envelope E protein and one in the nonstructural NS3 protein were found. The cytotoxicity of the mouse-neurovirulent DEN-1 variant was studied in different target cells in vitro and compared with the parental strain. FGA/NA d1d was more pathogenic for mouse neuroblastoma cells and attenuated for human hepatoma cells. Changes in virus replicative functions and virus assembly may account, in a large part, for the differences in the induction of apoptosis. Our data suggest that identified amino acid substitutions in the envelope E protein and viral RNA helicase NS3 may influence DEN-1 virus pathogenicity by altering viral growth.  
 Copyright 2000 Academic Press.

L19 ANSWER 8 OF 14 MEDLINE on STN  
 ACCESSION NUMBER: 2005316177 MEDLINE  
 DOCUMENT NUMBER: PubMed ID: 15855154  
 TITLE: Dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin (DC-SIGN)-mediated enhancement of **dengue** virus infection is independent of DC-SIGN internalization signals.  
 AUTHOR: Lozach Pierre-Yves; Burleigh Laura; Staropoli Isabelle; Navarro-Sanchez Erika; Harriague Julie; Virelizier Jean-Louis; Rey Felix A; **Despres Philippe**; Arenzana-Seisdedos Fernando; Amara Ali  
 CORPORATE SOURCE: Unite d'Immunologie Virale, Institut Pasteur Paris, 25-28, rue du Dr Roux, 75724 Paris Cedex 15, France.  
 SOURCE: The Journal of biological chemistry, (2005 Jun 24) Vol. 280, No. 25, pp. 23698-708. Electronic Publication: 2005-04-26. Journal code: 2985121R. ISSN: 0021-9258.

PUB. COUNTRY: United States  
 DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
 LANGUAGE: English  
 FILE SEGMENT: Priority Journals  
 ENTRY MONTH: 200509  
 ENTRY DATE: Entered STN: 21 Jun 2005  
 Last Updated on STN: 30 Sep 2005  
 Entered Medline: 29 Sep 2005

AB **Dengue** virus (DV) is a mosquito-borne flavivirus that causes hemorrhagic fever in humans. In the natural infection, DV is introduced into human skin by an infected mosquito vector where it is believed to target immature dendritic cells (DCs) and Langerhans cells (LCs). We found that DV productively infects DCs but not LCs. We show here that the interactions between DV E protein, the sole mannosylated glycoprotein present on DV particles, and the C-type lectin dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin (DC-SIGN) are essential for DV infection of DCs. Binding of mannosylated N-glycans on DV E protein to DC-SIGN triggers a rapid and efficient internalization of the viral glycoprotein. However, we observed that endocytosis-defective DC-SIGN molecules allow efficient DV replication, indicating that DC-SIGN endocytosis is dispensable for the internalization step in DV entry. Together, these results argue in favor of a mechanism by which DC-SIGN enhances DV entry and infection in cis. We propose that DC-SIGN concentrates mosquito-derived DV particles at the cell surface to allow efficient interaction with an as yet unidentified entry factor that is ultimately responsible for DV internalization and pH-dependent fusion into DCs.

L19 ANSWER 9 OF 14 MEDLINE on STN  
 ACCESSION NUMBER: 2000057956 MEDLINE  
 DOCUMENT NUMBER: PubMed ID: 10590151  
 TITLE: Alpha-glucosidase inhibitors reduce **dengue** virus production by affecting the initial steps of virion morphogenesis in the endoplasmic reticulum.  
 AUTHOR: Courageot M P; Frenkiel M P; Dos Santos C D; Deubel V; **Despres P**  
 CORPORATE SOURCE: Unite des Arbovirus et Virus des Fievres Hemorragiques, Institut Pasteur, 75724 Paris, France.  
 SOURCE: Journal of virology, (2000 Jan) Vol. 74, No. 1, pp. 564-72. Journal code: 0113724. ISSN: 0022-538X.  
 PUB. COUNTRY: United States  
 DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)  
 LANGUAGE: English  
 FILE SEGMENT: Priority Journals  
 ENTRY MONTH: 200001  
 ENTRY DATE: Entered STN: 24 Jan 2000  
 Last Updated on STN: 24 Jan 2000  
 Entered Medline: 10 Jan 2000

AB We report that endoplasmic reticulum alpha-glucosidase inhibitors have antiviral effects on **dengue** (DEN) virus. We found that glucosidase inhibition strongly affects productive folding pathways of the envelope glycoproteins prM (the intracellular glycosylated precursor of M [membrane protein]) and E (envelope protein): the proper folding of prM bearing unprocessed N-linked oligosaccharide is inefficient, and this causes delayed formation of prME heterodimer. The complexes formed between incompletely folded prM and E appear to be unstable, leading to a nonproductive pathway. Inhibition of alpha-glucosidase-mediated N-linked oligosaccharide trimming may thus prevent the assembly of DEN virus by affecting the early stages of envelope glycoprotein processing.

L19 ANSWER 10 OF 14 BIOSIS COPYRIGHT (c) 2006 The Thomson Corporation on STN

ACCESSION NUMBER: 2006:423308 BIOSIS  
DOCUMENT NUMBER: PREV200600423415  
TITLE: Comparative mechanistic studies of de novo RNA synthesis by flavivirus RNA-dependent RNA polymerases.  
AUTHOR(S): Selisko, Barbara; Dutartre, Helene; Guillemot, Jean-Claude; Debarnot, Claire; Benarroch, Delphine; Khromykh, Alexander; Despres, Philippe; Egloff, Marie-Pierre; Canard, Bruno [Reprint Author]  
CORPORATE SOURCE: CNRS, Case 925, 163 Ave Luminy, F-13288 Marseille 9, France Bruno.Canard@afmb.univ-mrs.fr  
SOURCE: Virology, (JUL 20 2006) Vol. 351, No. 1, pp. 145-158. CODEN: VIRLAX. ISSN: 0042-6822.  
DOCUMENT TYPE: Article  
LANGUAGE: English  
OTHER SOURCE: GenBank-AF481864; EMBL-AF481864; DDBJ-AF481864  
ENTRY DATE: Entered STN: 23 Aug 2006  
Last Updated on STN: 23 Aug 2006

AB Flavivirus protein NS5 harbors the RNA-dependent RNA polymerase (RdRp) activity. In contrast to the RdRps of hepaciviruses and pestiviruses, which belong to the same family of Flaviviridae, NS5 carries two activities, a methyltransferase (MTase) and a RdRp. RdRp domains of Dengue virus (DV) and West Nile virus (WNV) NS5 were purified in high yield relative to full-length NS5 and showed full RdRp activity. Steady-state enzymatic parameters were determined on homopolymeric template poly(rC). The presence of the MTase domain does not affect the RdRp activity. Flavivirus RdRp domains might bear more than one GTP binding site displaying positive cooperativity. The kinetics of RNA synthesis by four Flaviviridae RdRps were compared. In comparison to Hepatitis C RdRp, DV and WNV as well as Bovine Viral Diarrhea virus RdRps show less rate limitation by early steps of short-product formation. This suggests that they display a higher conformational flexibility upon the transition from initiation to elongation. (c) 2006 Elsevier Inc. All rights reserved.

L19 ANSWER 11 OF 14 EMBASE COPYRIGHT (c) 2006 Elsevier B.V. All rights reserved on STN

ACCESSION NUMBER: 2002062862 EMBASE  
TITLE: [Pathogenesis of dengue disease and apoptotic death].  
LA PATHOGENICITE DU VIRUS DE LA DENGUE ET LA MORT CELLULAIRE PAR APOPTOSE.  
AUTHOR: Courageot M.-P.; Despres P.  
CORPORATE SOURCE: P. Despres, U. Arbovirus et Virus Fievres Hemor., Institut Pasteur, 25, rue du Dr-Roux, 75724 Paris Cedex 15, France. pdespres@pasteur.fr  
SOURCE: Virologie, (2001) Vol. 5, No. 6, pp. 397-407. .  
Refs: 39  
ISSN: 1267-8694 CODEN: VIROFD  
COUNTRY: France  
DOCUMENT TYPE: Journal; General Review  
FILE SEGMENT: 004 Microbiology  
005 General Pathology and Pathological Anatomy  
LANGUAGE: French  
SUMMARY LANGUAGE: English; French  
ENTRY DATE: Entered STN: 1 Mar 2002  
Last Updated on STN: 1 Mar 2002

AB Dengue (DEN) is the most important vector-borne disease in tropical countries. DEN disease is caused by dengue virus, a

member of the flavivirus genus (family Flaviviridae). DEN is one major health concern in humans. DEN virus causes a spectrum of illnesses, ranging from a flu-like disease to DEN hemorrhagic fever, a fulminating illness that can progress to a shock syndrome and death. The pathogenesis of DEN disease is not well understood. Infection of target cells with DEN virus induces apoptosis. Changes in virus life cycle may account for differences in apoptosis induction. Determinants that may be relevant to DEN virus pathogenicity have been identified in the envelope E protein and viral helicase NS3. Intracellular synthesis of DEN envelope glycoproteins prM and E was sufficient to cause cell death. Induction of apoptosis may be linked to the presence of a pro-apoptotic **sequence** in the C-terminal region of prM.

L19 ANSWER 12 OF 14 WPIX COPYRIGHT 2006 THE THOMSON CORP on STN  
 ACCESSION NUMBER: 2005-761339 [78] WPIX  
 DOC. NO. CPI: C2005-232529  
 TITLE: Use of recombinant lentiviral vector for vaccination against infections by Flaviviridae, e.g. West Nile virus, **dengue**, yellow fever and hepatitis C.  
 DERWENT CLASS: B04 C06 D16  
 INVENTOR(S): CHARNEAU, P; **DESPRES, P**; FRENKIEL, M P; TANGY, F; FRENKIEL, M  
 PATENT ASSIGNEE(S): (CNRS) CNRS CENT NAT RECH SCI; (INSP) INST PASTEUR; (CNRS) CENT NAT RECH SCI  
 COUNTRY COUNT: 111  
 PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
FR 2870126	A1	20051118	(200578)*		61
WO 2005111221	A1	20051124	(200578)	EN	
RW: AT BE BG BW CH CY CZ DE DK EA EE ES FI FR GB GH GM GR HU IE IS IT KE LS LT LU MC MW MZ NA NL OA PL PT RO SD SE SI SK SL SZ TR TZ UG ZM ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE EG ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KM KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NA NG NI NO NZ OM PG PH PL PT RO RU SC SD SE SG SK SL SM SY TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW					

## APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
FR 2870126	A1	FR 2004-5366	20040517
WO 2005111221	A1	WO 2005-IB1753	20050516

PRIORITY APPLN. INFO: FR 2004-5366 20040517

AB FR 2870126 A UPAB: 20051205

NOVELTY - Use of a recombinant lentiviral vector (A) to prepare an immunogenic composition for prevention and/or treatment of infections by Flaviviridae, where (A) includes a polynucleotide fragment (I) that encodes at least one protein (II) from a virus of the family Flaviviridae or an immunogenic **peptide** (IIa), of at least 8 amino acids, from (II).

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(1) (A), As defined, that contains a polynucleotide that encodes at least one structural protein (or fragment) and optionally a non-structural

protein (or fragment);

- (2) cells, preferably eukaryotic, modified by (A);
- (3) method for producing proteins of Flaviviridae and/or their immunogenic fragments or viral pseudoparticles, by culturing cells of (2);
- (4) method of screening for antiviral compounds using the cells of (2);
- (5) method for diagnosing infection by Flaviviridae in a biological fluid by detecting antibody-antigen complex formation with the cells of (2) or pseudoparticles of (3); and
- (6) kit for methods (4) and (5) that contains the cells of (2).

ACTIVITY - Virucide; Hepatotropic; Antiinflammatory.

MECHANISM OF ACTION - Vaccine. Lentiviral vector TRIP Delta U3.CMV-Es(WNV), containing a 1.4 kb cDNA from West Nile virus (WNV), was used to immunize mice, at an intraperitoneal dose of 1 mu g. The anti-WNV antibody titer was 104 after 14 days and 2 plus or minus 105 after 23 days, with titer of antibodies that neutralize 90% of WNV loci of infections 10 and 20, respectively.

USE - (A) Are used to produce immunogenic compositions (vaccines) for treatment and/or prevention of Flaviviridae infections in humans and animals, particularly West Nile virus, **dengue**, yellow fever and hepatitis C.

Cells transformed with (A) are useful:

- (i) for preparing proteins, or their immunogenic fragments, from Flaviviridae;
- (ii) in screening for antiviral agents; and
- (iii) to diagnose Flaviviridae infections.

ADVANTAGE - (A) Induces a strong response; particularly it targets antigen-presenting cells and (I) becomes integrated into the genome, ensuring stable expression in vivo, especially in dendritic cells, eliminating the need for repeated administrations. (A) Is non-replicative; non-tumorigenic; not species restricted and does not require adjuvants.  
Dwg.0/5

L19 ANSWER 13 OF 14 WPIX COPYRIGHT 2006 THE THOMSON CORP on STN  
 ACCESSION NUMBER: 2003-058566 [05] WPIX  
 DOC. NO. NON-CPI: N2003-045379  
 DOC. NO. CPI: C2003-015061  
 TITLE: Identifying stimulators of oligoadenylate synthase family genes, useful as antiviral agents against Flavivirus, also mutated genes responsible for sensitivity to virus.  
 DERWENT CLASS: B04 D16 P14  
 INVENTOR(S): BONHOMME, F; **DESPRES, P**; DEUBEL, V; FRENKIEL, M  
 P; GUENET, J L; LUCAS, M; MASHIMO, T; MONTAGUTELLI, X;  
 SIMON, C D; FRENKIEL, M; GUENET, J; SIMON-CHAZOTTES, D  
 PATENT ASSIGNEE(S): (INSP) INST PASTEUR; (CNRS) CNRS CENT NAT RECH SCI  
 COUNTRY COUNT: 100  
 PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2002081741	A2	20021017	(200305)*	FR	93
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW					
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW					
FR 2823224	A1	20021011	(200305)		
AU 2002302677	A1	20021021	(200433)		

AU 2002302677 A8 20051020 (200615)

## APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2002081741	A2	WO 2002-FR1169	20020404
FR 2823224	A1	FR 2001-4598	20010404
AU 2002302677	A1	AU 2002-302677	20020404
AU 2002302677	A8	AU 2002-302677	20020404

## FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2002302677	A1 Based on	WO 2002081741
AU 2002302677	A8 Based on	WO 2002081741

PRIORITY APPLN. INFO: FR 2001-4598 20010404

AB WO 200281741 A UPAB: 20030121

NOVELTY - Identifying (M1) compounds (I) that can stimulate a gene (II) of the OAS (2'-5'-oligoadenylate synthase) family comprising:

(a) inducing expression of the OAS gene in a culture of cells from a non-human mammal (Flvr/Flvr or Flvr/Flvs; indicating resistance or sensitivity to Flavivirus infection);

(b) treating cells with test compound; and

(c) measuring activity of (II) relative to a control, is new.

DETAILED DESCRIPTION - Identifying (M1) compounds (I) that can stimulate a gene (II) of the OAS (2'-5'-oligoadenylate synthase) family comprising:

(a) inducing expression of the OAS gene in a culture of cells from a non-human mammal (Flvr/Flvr or Flvr/Flvs; indicating resistance or sensitivity to Flavivirus infection);

(b) treating cells with test compound; and

(c) measuring activity of (II) relative to a control, is new.

Step (a) involves addition of interferon (IFN) alpha or beta, or application of calcium stress, particularly addition of EGTA (ethylene glycol tetraacetic acid).

INDEPENDENT CLAIMS are also included for the following:

(1) genomic mammalian DNA (III), optionally human, corresponding to a locus for resistance to Flavivirus infection;

(2) evaluating (M2) sensitivity of an individual to Flavivirus infection and/or response to treatment with IFN;

(3) reagent (IV) for the new method containing specific primers or probes;

(4) transformed eukaryotic cells (V) containing (III), the corresponding cDNA or encoded proteins;

(5) non-human transgenic mammals (VI) that include at least one copy of (III) or the corresponding cDNA; and

(6) non-human transgenic mammals (VII) that contain at least one inactivated allele of an OAS gene.

ACTIVITY - Virucide; Hepatotropic; Antiinflammatory.

Primary rat neurons (from animals homozygous for the Flvs allele) were incubated for 2 hr in medium containing EGTA (causing a 350% increase in transcription of the OAS gene), then tested for infection by the neurovirulent IS-98-ST1 strain of West Nile virus. Treatment with EGTA caused a 50% reduction in the number of cells positive for viral antigen 24 hr after infection.

MECHANISM OF ACTION - Replacement or modulation of OAS activity, loss of which is associated with sensitivity to Flavivirus infection.

USE - Genomic nucleic acids, the corresponding cDNAs (also related vectors and cells containing such vectors), or their encoded proteins, are useful for screening for antivirals for treating Flavivirus infection (claimed). (I) are potentially useful as antiviral agents for treating infections by Flaviviruses (e.g. hepatitis C; **dengue**; yellow fever and varoious forms of encephalitis). Genomic OAS DNA and derived cDNA, also the encoded proteins, are useful:

- (a) for treating Flavivirus infection;
  - (b) in screening for anti-flavivirus agents, and
  - (c) for evaluating sensitivity of subjects to Flavivirus infection and their likely response to IFN treatment, e.g. to identify patients at risk of developing severe forms of such infections
- Dwg.0/17

L19 ANSWER 14 OF 14 WPIX COPYRIGHT 2006 THE THOMSON CORP on STN  
 ACCESSION NUMBER: 2001-418460 [45] WPIX  
 DOC. NO. NON-CPI: N2001-309994  
 DOC. NO. CPI: C2001-126641  
 TITLE: System for recombinant expression of flavivirus envelope proteins, useful in screening for potential antiviral agents, contains protein-encoding **sequences** and selection gene.  
 DERWENT CLASS: B04 D16 S03  
 INVENTOR(S): COURAGEOT, M; **DESPRES, P**; DEUBEL, V  
 PATENT ASSIGNEE(S): (INSP) INST PASTEUR  
 COUNTRY COUNT: 1  
 PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
CA 2290090	A1	20010607	(200145)*	FR	17

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
CA 2290090	A1	CA 1999-2290090	19991207

PRIORITY APPLN. INFO: CA 1999-2290090 19991207

AB CA 2290090 A UPAB: 20010813

NOVELTY - Inducible expression system (A), for mammalian cells (human or non-human) or insect cells, includes an insert comprising:

- (i) a **sequence** that encodes at least one flavivirus envelope protein (FEP);
- (ii) a selection gene; and
- (iii) optionally a gene that encodes a cellular hormone receptor.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

- (1) selection methods using (A); and
- (2) inhibitors (IV) of flavivirus (FV), able to inhibit association of FEP produced in (A).

ACTIVITY - Antiviral.

MECHANISM OF ACTION - Modulating folding of, and interactions between, FEP, particularly by inhibition of alpha -glucosidase which is essential for protein maturation before virion assembly.

USE - (A) is used:

- (i) to test efficiency of potential inhibitors directed against FEP;
- (ii) to test folding and efficiency of association of FEP;
- (iii) to validate inhibitory activity of drugs;

(iv) for in vivo or in vitro evaluation of the protective capacity of agents that prevent cytotoxicity of FEP; and

(v) to identify, in vivo or in vitro, cellular genes that are activated in response to FEP.

Inhibitors identified using (A) are potentially useful for treating flavivirus infections, particularly the various forms of **dengue** fever, yellow fever, hepatitis C and G, or Japanese/tick-borne encephalitis.

Dwg.0/0

=>



GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:29:54 ; Search time 107.75 Seconds  
(without alignments)  
38.190 Million cell updates/sec

Title: DENGUE\_SEROTYPE1

Perfect score: 9

Sequence: 1 vetflrhp 9

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2589679 seqs, 457216429 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2571072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : A\_Geneseq\_8.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*
- 10: Geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	88.9	39	9	Adw12582 M1-40/DEN
2	8	88.9	48	9	Adw12588 p(95-114)
3	6	66.7	278	8	Adq25888 Human GPC
4	6	66.7	826	5	Abb07253 Human nov
5	6	66.7	827	6	Abu07568 Human sec
6	6	66.7	904	4	Abg09947 Novel hum
7	6	66.7	924	5	Aab71323 Human GCR
8	6	66.7	953	7	Ade34415 Human G-p
9	6	66.7	994	5	Abb07252 Human nov
10	6	66.7	994	5	Aau99808 Novel hum
11	6	66.7	994	7	Ade34425 Human G-p
12	6	66.7	994	8	Ado28977 Human nov
13	6	66.7	994	8	Adq25892 Human nov
14	6	66.7	1018	5	Aae25061 Human G-p
15	6	66.7	1070	6	Abu07567 Human sec
16	6	66.7	1131	4	Abg11655 Novel hum
17	6	66.7	1232	7	Adf70474 Orphan re
18	5	55.6	28	10	Aee37134 Human ser
19	5	55.6	34	4	Aau17769 Novel hum
20	5	55.6	34	7	Adg41149 Human res
21	5	55.6	34	7	Adi96923 Human res
22	5	55.6	52	6	Abg99963 Human nov
23	5	55.6	60	4	Aau42843 Propionib

Abm393362	Propionib	60	6	ABM393362
Aam86926	Human imm	62	4	AAM86926
Aag98737	Human cel	55.6	64	AAG98737
Aau50032	Propionib	55.6	64	Aau50032
Abm46551	Propionib	55.6	64	ABM46551
Abp03674	Human ORF	55.6	73	ABP03674
Aag98736	Human cel	55.6	75	AAG98736
Aam99844	Human exc	55.6	76	AAM99844
Aam42659	Human kid	55.6	76	AAM42659
Abp34862	Human ORF	55.6	82	ABP34862
Adt58131	Plant pol	55.6	85	ADT58131
Aau51518	Propionib	55.6	86	Aau51518
Abm48037	Propionib	55.6	86	ABM48037
Adx94950	Plant ful	55.6	89	ADX94950
Adk36992	Novel hum	55.6	90	ADK36992
Abm94143	M. xanthu	55.6	98	ABM94143
Adc14235	Human enz	55.6	102	ADC14235
Aay74113	Human pro	55.6	104	AAY74113
Aau09103	Novel hum	55.6	106	Aau09103
Aao07214	Human pol	55.6	108	AAO07214
Abu51319	Helicobac	55.6	110	ABU51319
Abp75900	Human sec	55.6	111	ABP75900
Adx90486	Plant ful	55.6	124	ADX90486
Aea79622	IC6 MAB h	55.6	125	AEA79622
Aec39351	Human IC6	55.6	125	AEC39351
Aam50235	Catalpa s	55.6	127	AAM50235
Aau76417	Catalpa l	55.6	127	Aau76417
Adk36828	Novel hum	55.6	146	ADK36828
Ady23809	Plant ful	55.6	149	ADY23809
Aaw07588	Fibroblas	55.6	154	Aaw07588
Aaw07587	Fibroblas	55.6	154	Aaw07587
Aaw07589	Fibroblas	55.6	154	Aaw07589
Aaw07590	Fibroblas	55.6	154	Aaw07590
Aay90462	Mutant hu	55.6	154	AAY90462
Aay90464	Mutant hu	55.6	154	AAY90464
Aay90460	Saporin e	55.6	154	AAY90460
Aay90461	Mutant hu	55.6	154	AAY90461
Aay90463	Mutant hu	55.6	154	AAY90463
Adc34717	Human fib	55.6	154	ADC34717
Adc34715	Human fib	55.6	154	ADC34715
Adc34718	Human fib	55.6	154	ADC34718
Adc34716	Human fib	55.6	154	ADC34716
Adh92147	Fibroblas	55.6	154	ADH92147
Adh92148	Fibroblas	55.6	154	ADH92148
Adh92145	Codon opt	55.6	154	ADH92145
Adh92146	Fibroblas	55.6	154	ADH92146
Abu51724	Helicobac	55.6	158	ABU51724
Aau64446	Propionib	55.6	175	Aau64446
Abm60965	Propionib	55.6	175	ABM60965
Aae06639	Human alp	55.6	179	AAE06639
Abm92796	M. xanthu	55.6	188	ABM92796
Aam93280	Human pol	55.6	189	AAM93280
Adl30724	Human pro	55.6	189	ADL30724
Abm58403	Human NOV	55.6	198	ABM58403
Abb90287	Human pol	55.6	201	ABB90287
Aed00443	Lactobaci	55.6	215	AED00443
Aam39930	Human pol	55.6	216	AAM39930
Aed00317	Lactobaci	55.6	225	AED00317
Aau29056	Human PRO	55.6	234	Aau29056
Aam39929	Human pol	55.6	234	AAM39929
Aab87532	Human PRO	55.6	234	AAB87532
Abg95857	Human sec	55.6	234	ABG95857
Abb84847	Human PRO	55.6	234	ABB84847
Abm95453	Human ang	55.6	234	ABM95453
Abu58432	Human PRO	55.6	234	ABU58432
Abu87980	Novel hum	55.6	234	ABU87980
Abu84295	Human sec	55.6	234	ABU84295
Abm66169	Human sec	55.6	234	ABM66169
Abm65559	Human sec	55.6	234	ABM65559
Abu99499	Human sec	55.6	234	ABU99499
Abu82738	Human PRO	55.6	234	ABU82738
Abu89859	Novel hum	55.6	234	ABU89859
Abm68108	Human sec	55.6	234	ABM68108



XX 30-JUN-2003; 2003US-00608029.  
XX (INSP ) INST PASTEUR.  
XX Despres P, Catteau A;  
XX WPI; 2005-047647/05.  
XX N-PSDB; ADW12589.  
XX New isolated and purified Apoptom peptide comprises 9 amino acids, useful  
XX as a vaccine for preventing or treating pathological conditions from non-  
XX specific febrile illnesses to severe hemorrhagic manifestations or  
XX encephalitic syndromes.  
XX Disclosure; SEQ ID NO 35; 30pp; English.  
XX  
XX The present invention relates to an isolated and purified Apoptom  
XX peptide. The invention is useful as a vaccine for the prevention and  
XX treatment of pathological conditions from non-specific febrile illnesses  
XX to severe hemorrhagic manifestations, encephalitic syndromes and these  
XX pathological conditions are linked to Flavivirus infection or cancers.  
XX The invention is also useful in gene therapy. The present sequence is a  
XX p(95-114) EGFP (enhanced green fluorescent protein) (M1-M40)DEN (dengue)-2  
XX (136F) plasmid DNA encoded protein.  
XX  
XX Sequence 48 AA;  
SQ  
Query Match 88.9%; Score 8; DB 9; Length 48;  
Best Local Similarity 100.0%; Pred. No. 0.044;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 ETWFLRHP 9  
DB 41 ETWFLRHP 48  
|||||  
RESULT 3  
ADQ25888  
ID ADQ25888 standard; protein; 278 AA.  
XX  
XX ADQ25888;  
XX  
XX 23-SEP-2004 (first entry)  
XX Human GPCR related protein #1.  
XX  
XX receptor; GPCR; guanosine triphosphate-binding protein-coupled receptor;  
XX human.  
XX Homo sapiens.  
XX  
XX WO2004055186-A1.  
XX  
XX 01-JUL-2004.  
XX  
XX 18-DEC-2003; 2003WO-JP016245.  
XX  
XX 18-DEC-2002; 2002JP-00366417.  
XX  
XX 03-MAR-2003; 2003JP-00055691.  
XX  
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.  
XX  
XX Suwa M, Asai K, Akiyama Y, Aburatani H;  
XX  
XX WPI; 2004-500216/47.  
XX  
XX N-PSDB; ADQ25887.  
XX  
XX New polynucleotide encoding guanosine triphosphate-binding protein-  
XX coupled receptor, for use in developing a therapeutic agent for medical  
XX treatment.  
XX

PS Example 5; SEQ ID NO 16; 104pp; Japanese.  
XX  
XX The present invention provides the protein and coding sequences of a  
XX human guanosine triphosphate-binding protein-coupled receptor (GPCR). The  
XX sequences are useful for treating diseases related to the abnormality of  
XX the expression of GPCR, and for developing a therapeutic agent for  
XX medical treatment. The present sequence is a protein shown in the  
XX exemplification of the invention.  
XX  
XX Sequence 278 AA;  
SQ  
Query Match 66.7%; Score 6; DB 8; Length 278;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 WFLRHP 9  
DB 185 WFLRHP 190  
|||||  
RESULT 4  
ABB07253  
ID ABB07253 standard; protein; 826 AA.  
XX  
XX ABB07253;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Human novel GPCR (NGPCR) protein.  
XX  
XX G coupled protein receptor; GPCR; NGPCR; cytostatic; anorectic; cancer;  
XX antiinflammatory; immunosuppressive; antidiabetic; human.  
XX  
XX Homo sapiens.  
XX  
XX WO200187932-A2.  
XX  
XX 22-NOV-2001.  
XX  
XX 11-MAY-2001; 2001WO-US015048.  
XX  
XX 12-MAY-2000; 2000US-0203875P.  
XX  
XX 30-MAY-2000; 2000US-0207932P.  
XX  
XX (LEXI-) LEXICON GENETICS INC.  
XX  
XX Hu Y, Nepomnichy B, Wang X, Walke DW, Gerhardt B, Turner CA;  
XX  
XX WPI; 2002-114231/15.  
XX  
XX N-PSDB; ABA94352.  
XX  
XX New polypeptide, useful for generation of antibodies and for screening  
XX compounds for treatment of mental, biological or medical disorders and  
XX diseases, comprises the isolated G coupled protein receptor polypeptide.  
XX  
XX Claim 8; Page 81-83; 85pp; English.  
XX  
XX The invention provides novel G coupled protein receptor (GPCR) proteins  
XX and polynucleotides encoding the same. The novel GPCR (NGPCR) proteins  
XX can be expressed by standard recombinant methodology. The NGPCR proteins  
XX and polynucleotides are useful for diagnosis, in treatment of diseases,  
XX drug screening, clinical trial monitoring, for treatment of physiological  
XX or behavioural disorders, for the detection of mutant GPCRs or  
XX inappropriately expressed GPCR for the diagnosis of disease, and for  
XX screening drugs effective in the treatment of the symptomatic or  
XX phenotypic manifestations of perturbing the normal function of GPCR in  
XX the body. The NGPCR proteins are useful for the generation of antibodies,  
XX as reagents in diagnostic assays, for the identification of other  
XX cellular gene products related to a GPCR, as reagents in assays for  
XX screening compounds that can be used as pharmaceutical reagents for the  
XX therapeutic treatment of mental, biological or medical disorders and  
XX diseases, and for identifying compounds useful in the therapeutic  
XX treatment of obesity, inflammation, immune disorders, diabetes, heart and

CC coronary disease, metabolic disorders, and cancer. The present sequence  
 CC represents a human NPCR protein  
 XX  
 SQ Sequence 826 AA;  
 Query Match 66.7%; Score 6; DB 5; Length 826;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRHP 9  
 Db 21 WFLRHP 26  
 RESULT 5  
 ABU07568  
 ID ABU07568 standard; protein; 827 AA.  
 XX  
 AC ABU07568;  
 DT 20-MAR-2003 (first entry)  
 XX  
 DE Human secretin type G protein-coupled receptor #2.  
 XX  
 KW Human; receptor; GPCR; G protein-coupled receptor; secretin; obesity;  
 KW cardiovascular disorder; diabetes; infection; HIV; pain; cancer;  
 KW human immunodeficiency virus infection; anorexia; bulimia; asthma;  
 KW Parkinson's disease; acute heart failure; hypotension; hypertension;  
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;  
 KW ulcer; allergy; benign prostatic hypertrophy; psychosis;  
 KW neurological disorder; anxiety; schizophrenia; manic depression;  
 KW delirium; dementia; mental retardation; dyskinesia; Huntington's disease;  
 KW Tourette's syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 827 /note= "Encoded by GA"  
 FT  
 XX  
 PN WO200299106-A2.  
 PD 12-DEC-2002.  
 XX  
 PF 03-JUN-2002; 2002WO-EP006031.  
 XX  
 PR 04-JUN-2001; 2001US-0294998P.  
 PR 26-JUL-2001; 2001US-0307608P.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Koehler RH, Smolyar A;  
 DR WPI; 2003-140623/13.  
 DR N-PSDB; ABX15279.  
 XX  
 PT New isolated polynucleotide encoding human secretin-type G protein-  
 PT coupled receptor (GPCR) polypeptides, useful for preventing or treating  
 PT diseases associated with GPCR dysfunction, e.g. cardiovascular disease or  
 PT diabetes.  
 XX  
 PS Claim 1; Fig 7; 127pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide which: (a) encodes a  
 CC human secretin-type G protein-coupled receptor (GPCR) polypeptide; (b)  
 CC comprises a sequence appearing as ABX15278 and ABX15279; (c) hybridises  
 CC under stringent conditions to the polynucleotide in (A) and (B); (d) has  
 CC a sequence deviating from (A)-(C) due to the degeneration of the genetic  
 CC code, or represents a fragment, derivative or allelic variation of (A)-  
 CC (D). Also included are an expression vector containing the above  
 CC polynucleotide, a host cell containing the expression vector, a  
 CC substantially purified human secretin-type GPCR polypeptide, methods of  
 CC screening for agents which modulate or decrease the activity of a human

CC secretin-type GPCR, methods of reducing the activity of the human  
 CC secretin-type GPCR, the identified modulators. The polynucleotide is  
 CC useful in preventing, ameliorating, or treating diseases associated with  
 CC human secretin-type GPCR dysfunction. The polynucleotide may also be used  
 CC as hybridisation probes or primers, and in diagnostic assays or in  
 CC genetic testing. The methods are useful in producing and detecting the  
 CC polynucleotide and polypeptide and in screening for agents that modulate  
 CC the activity of the human secretin-type GPCR. The expression vector or  
 CC the reagent is useful in preparing a medicament for modulating the  
 CC activity of a human secretin-type GPCR in a disease, such as a  
 CC cardiovascular disorder, obesity, diabetes, infections (bacterial, viral,  
 CC fungal and protozoan), HIV (human immunodeficiency virus) infection,  
 CC pain, cancer, anorexia, bulimia, asthma, Parkinson's disease, acute heart  
 CC failure, hypotension, hypertension, urinary retention, osteoporosis,  
 CC angina pectoris, myocardial infarction, ulcers, allergies, benign  
 CC prostatic hypertrophy, psychosis, neurological disorders (e.g. anxiety,  
 CC schizophrenia, manic depression, delirium, dementia, mental retardation,  
 CC dyskinesias, Huntington's disease and Tourette's syndrome). The present  
 CC sequence represents a human secretin type GPCR of the invention  
 XX  
 SQ Sequence 827 AA;  
 Query Match 66.7%; Score 6; DB 6; Length 827;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRHP 9  
 Db 21 WFLRHP 26  
 RESULT 6  
 ABG09947  
 ID ABG09947 standard; protein; 904 AA.  
 XX  
 AC ABG09947;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #9938.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS74134.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PS Claim 20; SEQ ID NO 40306; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (II) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (III). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 904 AA;

Query Match 66.7%; Score 6; DB 4; Length 904;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9  
 Db 156 WFLRHP 161  
 |||||

RESULT 7  
 ID AAB71323 standard; protein; 924 AA.

AC AAB71323;

DT 19-NOV-2002 (first entry)

DE Human GCREC-2 INCYTE ID 7474890CD1 SEQ ID 2.

CC GCREC; Human; G-protein coupled receptor; anti-HIV; antiarteriosclerotic;  
 KW cytosolic; neuroprotective; antiparkinsonian; hepatotropic; laxative;  
 KW cerebroprotective; antiinflammatory; virucide; antibacterial; fungicide;  
 KW protozoacide; cirrhosis; cancer; stroke; Alzheimer's disease; AIDS;  
 KW Parkinson's disease; Crohn's disease; constipation; infection; receptor;  
 KW gene therapy.

OS Homo sapiens.

XX WO200263004-A2.

PD 15-AUG-2002.

PF 06-FEB-2002; 2002WO-US003635.

XX 07-FEB-2001; 2001US-0267322P.

PR 23-FEB-2001; 2001US-0271215P.

PR 08-MAR-2001; 2001US-0274551P.

PR 23-MAR-2001; 2001US-0278507P.

PR 30-MAR-2001; 2001US-0280597P.

PR 02-APR-2001; 2001US-0281107P.

PR 06-APR-2001; 2001US-0282121P.

XX (INCY-) INCYTE GENOMICS INC.

PA Baughn MR, Tribouley CM, Nguyen DB, Thornton M, Yao MG;

PI Kalliock DA, Gandhi AR, Wallia NK, Arvizu C, Elliott VS, Hafalia AJA;

PI Ramkumar J, Pei J, Tang YT, Yue H, Reddy R, Butford N, Lu DAM;

PI Graul RC, Khan FA, Walsh RT, Ison CH, Richardson TW, Griffin JA;

PI Warren BA, Yang J, Lee EA, Harland L;

PT New human G-protein coupled receptors (GCREC), useful for diagnosing or  
 PT treating a disease or condition associated with decreased expression or  
 PT over expression of functional GCRECs e.g. cancer, Alzheimer's and  
 PT Parkinson's.

XX Claim 63; Page 160-163; 239pp; English.

XX This invention describes novel polypeptides which have anti-HIV,  
 CC antiarteriosclerotic, cytosolic, neuroprotective, antiparkinsonian,  
 CC hepatotropic, laxative, cerebroprotective, antiinflammatory, virucide,  
 CC antibacterial, fungicide and protozoacide activity. The products of the  
 CC invention are useful for treating a disease or condition associated with  
 CC decreased expression or over expression of functional G-protein coupled  
 CC receptors (GCREC), while antibodies generated against the polypeptide of  
 CC the invention are useful for diagnosing a condition or disease associated  
 CC with the expression of GCREC e.g. arteriosclerosis, cirrhosis, cancer,  
 CC stroke, Alzheimer's disease, Parkinson's disease, Crohn's disease,  
 CC constipation, AIDS, or bacterial, viral, fungal or protozoal infections.  
 CC The compounds described in the invention can be used for gene therapy.  
 CC AAB71322-AAB71369 represent the GCREC proteins encoded by AAF88580-  
 CC AAF88627 described in the disclosure of the invention

XX Sequence 924 AA;

Query Match 66.7%; Score 6; DB 5; Length 924;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9  
 Db 141 WFLRHP 146  
 |||||

RESULT 8  
 ADE34415

ID ADE34415 standard; protein; 953 AA.

XX ADE34415;

XX 29-JAN-2004 (first entry)

DT Human G-protein coupled receptor protein #SEQ ID 35.

CC Cytostatic; antiinflammatory; hepatotropic; nephrotropic; dermatological;  
 KW antiarthritic; antiaesthetic; antidiabetic; hypotensive; antitumor;  
 KW antilipemic; antiarteriosclerotic; neurotropic; neuroprotective; anorectic;  
 KW immunomodulator; uropathic; antiinfertility; G-protein coupled receptor;  
 KW GPCR; GPCR185; GPCR186; GPCR187; GPCR188; GPCR189; GPCR222; GPCR223;  
 KW hepatitis; nephritis; dermatitis; pancreatitis; rheumatoid arthritis;  
 KW osteoarthritis; atopic dermatitis; asthma; diabetes; hypertension;  
 KW inflammatory bowel disease; gastric ulcer; arteriosclerosis;  
 KW hyperlipemia; Alzheimer's disease; dementia; obesity; pulmonary fibrosis;  
 KW renal fibrosis; immune deficiency; infertility; urinary blockage; cancer.

XX Homo sapiens.

OS WO2003078632-A1.

PN 25-SEP-2003.

PD 14-MAR-2003; 2003WO-JP003050.

PF 15-MAR-2002; 2002JP-00071567.

PR 14-MAY-2002; 2002JP-00138013.

PR 28-FEB-2003; 2003JP-00054663.

XX (NIBS) JAPAN TOBACCO INC.

PI Watanabe H, Nozaki Y;

XX WPI; 2003-722435/68.

PT G-protein coupled receptor proteins, genes encoding them and antibodies

PT recognizing them for treatment and diagnosis of cancer, inflammatory and  
 PT gastrointestinal disorders.  
 XX Example; SEQ ID NO 35; 274pp; Japanese.  
 XX  
 CC The invention relates to G-protein coupled receptor proteins of human  
 CC origin. These proteins include GPCR185, GPCR186, GPCR187, GPCR188,  
 CC GPCR189, GPCR222 and GPCR223. Proteins of the invention are used in the  
 CC treatment and prevention of diseases associated with inflammation,  
 CC angiogenesis and tissue neogenesis, including hepatitis, nephritis,  
 CC dermatitis, pancreatitis, rheumatoid arthritis, osteoarthritis, atopic  
 CC dermatitis, asthma, diabetes, hypertension, inflammatory bowel disease,  
 CC gastric ulcer, arteriosclerosis, hyperlipemia, Alzheimer's disease,  
 CC dementia, obesity, pulmonary fibrosis, renal fibrosis, immune deficiency,  
 CC infertility, urinary blockage and cancer (such as cancer of the brain,  
 CC neck, tongue, lung, breast, pancreas, stomach, colon, duodenum, prostate,  
 CC bladder, ovary, womb or rectum). Primers of the invention are devised and  
 CC synthesised based on G-protein coupled receptor consensus sequences and  
 CC used for 5'-RACE (rapid amplification of cDNA ends) and 3'-RACE  
 CC amplification of human cDNA derived from adrenal and visual cortex RNA.  
 CC Sequences given in ABE34534-ADE34533 represent human G-protein coupled  
 CC receptor proteins, genes encoding them, and primers for the amplification  
 CC of these sequences.  
 XX  
 SQ Sequence 953 AA;

Query Match 66.7%; Score 6; DB 7; Length 953;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9  
 |||||  
 Db 141 WFLRHP 146

RESULT 9  
 ABB07252  
 ID ABB07252 standard; protein; 994 AA.  
 XX  
 AC ABB07252;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human novel GPCR (NGPCR) protein.  
 XX  
 KW G coupled protein receptor; GPCR; NGPCR; cytostatic; anorectic; cancer;  
 KW antiinflammatory; immunosuppressive; antidiabetic; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200187932-A2.  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 11-MAY-2001; 2001WO-US015048.  
 XX  
 PR 12-MAY-2000; 2000US-0203875P.  
 PR 30-MAY-2000; 2000US-0207932P.  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Hu Y, Nepomnichy B, Wang X, Walke DW, Gerhardt B, Turner CA;  
 XX  
 DR WPI; 2002-114231/15.  
 DR N-PSDB; ABA94351.  
 XX  
 PT New polypeptide, useful for generation of antibodies and for screening  
 PT compounds for treatment of mental, biological or medical disorders and  
 PT diseases, comprises the isolated G coupled protein receptor polypeptide.  
 XX  
 PS Claim 8; Page 78-80; 85pp; English.  
 XX  
 CC The invention provides novel G coupled protein receptor (GPCR) proteins

CC and polynucleotides encoding the same. The novel GPCR (NGPCR) proteins  
 CC can be expressed by standard recombinant methodology. The NGPCR proteins  
 CC and polynucleotides are useful for diagnosis, in treatment of diseases,  
 CC drug screening, clinical trial monitoring, for treatment of physiological  
 CC or behavioural disorders, for the detection of mutant GPCRs or  
 CC inappropriately expressed GPCR for the diagnosis of disease, and for  
 CC screening drugs effective in the treatment of the symptomatic or  
 CC phenotypic manifestations of perturbing the normal function of GPCR in  
 CC the body. The NGPCR proteins are useful for the generation of antibodies,  
 CC as reagents in diagnostic assays, for the identification of other  
 CC cellular gene products related to a GPCR, as reagents in assays for  
 CC screening compounds that can be used as pharmaceutical reagents for the  
 CC therapeutic treatment of mental, biological or medical disorders and  
 CC diseases, and for identifying compounds useful in the therapeutic  
 CC treatment of obesity, inflammation, immune disorders, diabetes, heart and  
 CC coronary disease, metabolic disorders, and cancer. The present sequence  
 CC represents a human NGPCR protein  
 XX  
 SQ Sequence 994 AA;

Query Match 66.7%; Score 6; DB 5; Length 994;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9  
 |||||  
 Db 189 WFLRHP 194

RESULT 10  
 AAU99808  
 ID AAU99808 standard; protein; 994 AA.  
 XX  
 AC AAU99808;  
 XX  
 DT 07-OCT-2002 (first entry)  
 XX  
 DE Novel human G protein-coupled receptor hTGR21-1.  
 XX  
 KW Human; G protein-coupled; receptor; hTGR21; central nervous disease;  
 KW endocrine disease; metabolic disease; cancer; inflammation; nootropic;  
 KW circulatory disorder; respiratory disorder; digestive disorder;  
 KW immune system disorder; infection; gene therapy; neuroprotective;  
 KW antiinflammatory; immunomodulator; cardiant; antimicrobial; cytostatic;  
 KW gene therapy; hTGR21-1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200253593-A1.  
 XX  
 PD 11-JUL-2002.  
 XX  
 PF 27-DEC-2001; 2001WO-JP011530.  
 XX  
 PR 28-DEC-2000; 2000JP-00400625.  
 PR 13-APR-2001; 2001JP-00115916.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Miwa M, Ito T, Shintani Y, Miyajima N;  
 XX  
 DR WPI; 2002-528854/56.  
 DR N-PSDB; ABK88069.  
 XX  
 PT Human kidney-originated G protein-coupled receptor protein hTGR21 and  
 PT encoding DNA, for developing drugs to treat e.g. central nervous  
 PT diseases, endocrine diseases, inflammations and diseases of digestive  
 PT system.  
 XX  
 PS Claim 1; Page 106-110; 143pp; Japanese.  
 XX  
 CC The invention describes a novel human kidney-originated G protein-coupled  
 CC receptor protein hTGR21 and the DNA encoding it. The proteins, DNAs and

CC	infertility, urinary blockage and cancer (such as cancer of the brain,
CC	neck, tongue, lung, breast, pancreas, stomach, colon, duodenum, prostate,
CC	bladder, ovary, womb or rectum) . Primers of the invention are devised and
CC	synthesised based on G-protein coupled receptor consensus sequences and
CC	used for 5'-RACE (rapid amplification of cDNA ends) and 3'-RACE
CC	amplification of human cDNA derived from adrenal and visual cortex RNA.
CC	Sequences given in ADE34534-ADE34533 represent human G-protein coupled
CC	receptor proteins, genes encoding them, and primers for the amplification
CC	of these sequences.
XX	
XX	
SQ	Sequence 994 AA;
	Query Match 66.7%; Score 6; DB 7; Length 994;
	Best Local Similarity 100.0%; Pred. No. 1.3e+02;
	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	4 WFLRHP 9
Db	189 WFLRHP 194
RESULT 12	
AD028977	
ID	AD028977 standard; protein; 994 AA.
XX	
XX	AD028977;
XX	
XX	29-JUL-2004 (first entry)
XX	
XX	Human novel GPCR PGR23, SEQ ID NO:76.
DE	
XX	G protein-coupled receptor; GPCR; drug screening; diagnosis;
XX	transgenic mouse; neurological disorder; adrenal gland disorder;
KW	colon disorder; intestinal disorder; cardiovascular disorder;
KW	muscular disorder; blood disorder; immune disorder; bone disorder;
KW	joint disorder; metabolic disorder; nutritive disorder; cancer;
KW	kidney disorder; liver disorder; lung disorder; breast disorder;
KW	ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW	skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW	thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW	cytostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW	CNS; central nervous system; respiratory; antiarrhythmic; antiadrenergic;
KW	virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW	dermatological; antitumor; antithyroid; antiallergic; anorectic;
KW	immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW	receptor.
XX	
XX	Homo sapiens.
OS	
XX	WO2004040000-A2.
XX	
XX	13-MAY-2004.
XX	
XX	09-SEP-2003; 2003WO-US028226.
XX	
XX	09-SEP-2002; 2002US-0409303P.
XX	
XX	09-APR-2003; 2003US-0461329P.
XX	
XX	(PRIM-) PRIMAL INC.
PA	
XX	Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI	Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
PI	
DR	WPI: 2004-390329/36.
DR	N-PSDB; AD028978.
XX	
XX	Novel mammalian G protein coupled receptors, useful for identifying
PT	compounds that modulates diagnosing and treating disease condition
PT	associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT	pectoris, Parkinson's disease.
XX	
XX	Claim 1; SEQ ID NO 76; 542pp; English.
PS	
XX	

CC The invention relates to human and mouse G protein-coupled receptors  
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to  
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids  
 CC of the invention; methods of treating, preventing or diagnosing diseases  
 CC associated with GPCRs of the invention; methods of screening for  
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic  
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a  
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells  
 CC from the transgenic mice; kits comprising several mice, each of which has  
 CC a mutation in a different GPCR gene of the invention; and kits comprising  
 CC probes which hybridise to GPCR polynucleotides of the invention. The  
 CC invention further discloses variants of the GPCR polypeptides and vectors  
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may  
 CC be used in the diagnosis, treatment or prevention of a wide variety of  
 CC diseases including neurological disorders (e.g., Alzheimer's disease,  
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);  
 CC disorders of the adrenal gland; disorders of the colon or intestine  
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel  
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or  
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,  
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or  
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid  
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,  
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related  
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,  
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and  
 CC thyroid (e.g., cancers). The present sequence represents a GPCR of the  
 CC invention. Note: The full sequence data for this patent did not form part  
 CC of the printed specification; those sequences not shown were obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 994 AA;

Query Match 66.7%; Score 6; DB 8; Length 994;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9  
 DB 189 WFLRHP 194  
 |||||

RESULT 13  
 ADQ25892 standard; protein; 994 AA.

XX ADQ25892;  
 XX 23-SEP-2004 (first entry)

DE Human guanosine triphosphate-binding protein-coupled receptor.

XX receptor; GPCR; guanosine triphosphate-binding protein-coupled receptor;  
 XX human.

XX Homo sapiens.

XX WO2004055186-A1.

XX 01-JUL-2004.

XX 18-DEC-2003; 2003WO-JP016245.

XX 18-DEC-2002; 2002JP-00366417.

PR 03-MAR-2003; 2003JP-00055691.

XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

XX Suwa M, Asai K, Akiyama Y, Aburatani H;

XX WPI; 2004-500216/47.

DR N-PSDB; ADQ25891.  
 XX New polynucleotide encoding guanosine triphosphate-binding protein-  
 PT coupled receptor, for use in developing a therapeutic agent for medical  
 PT treatment.

PS Claim 1; SEQ ID NO 20; 104pp; Japanese.

XX The present invention provides the protein and coding sequences of a  
 CC human guanosine triphosphate-binding protein-coupled receptor (GPCR). The  
 CC sequences are useful for treating diseases related to the abnormality of  
 CC the expression of GPCR, and for developing a therapeutic agent for  
 CC medical treatment. The present sequence is the protein of the invention.

XX Sequence 994 AA;

Query Match 66.7%; Score 6; DB 8; Length 994;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9  
 DB 189 WFLRHP 194  
 |||||

RESULT 14

AE25061 standard; protein; 1018 AA.

XX AE25061;

XX 30-OCT-2002 (first entry)

DE Human G-protein coupled receptor (GCRC)-1 protein.

XX Human; G-protein coupled receptor; GCRC; olfactory; taste sensation;  
 KW cell proliferative disorder; actinic keratosis; leukaemia; metabolic;  
 KW epilepsy; Alzheimer's disease; cardiovascular; hypertension; vitreous;  
 KW angina pectoris; myocardial infarction; gastrointestinal; anorexia;  
 KW cholecystitis; Crohn's disease; inflammatory; hypotensive; cardiac;  
 KW acquired immune deficiency syndrome; anaemia; asthma; hepatotropic;  
 KW diabetes; obesity; infection; transgenic; gene therapy; cytostatic;  
 KW anticonvulsant; neuroprotective; antiinflammatory; neurologic;  
 KW nootropic; anorectic; autoimmune; receptor.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT Peptide /label= Signal\_peptide  
 FT Peptide /label= Signal\_peptide

FT Protein /note= "Human mature GCRC-1 protein"  
 FT Protein /note= "Human mature GCRC-1 protein"

FT Domain /note= "Latrophilin/CL-1-like GPS domain"  
 XX WO200246230-A2.

XX 13-JUN-2002.

XX 05-DEC-2001; 2001WO-US046659.

XX 08-DEC-2000; 2000US-0254323P.

PR 21-DEC-2000; 2000US-0255564P.

PR 19-JAN-2001; 2001US-0262848P.

XX (INCY-) INCYTE GENOMICS INC.

XX Kallick DA, Baughn MR, Lu DAM, Yue H, Graul RC, Lu Y, Ding L;



```

PI XX Tribouley CM, Tang YT, Gandhi AR, Thornton M;
DR XX WPI; 2002-519657/55.
DR XX N-PSDB; AAD40625.
XX
XX Novel isolated human G-protein coupled receptor protein useful for
PT diagnosing, treating, preventing hypertension, myocardial infarction,
PT anorexia, cholecystitis, anaemia, asthma, diabetes, obesity, Alzheimer's
PT disease.
XX
XX Claim 1; Page 117-120; 136pp; English.
XX
XX The invention relates to human G-protein coupled receptors (GCREC) and
CC their corresponding nucleic acids. GCREC is useful in screening for
CC compounds which acts as its agonist or antagonist and is also useful for
CC preparing a polyclonal or monoclonal antibody. GCREC is useful for
CC identifying a compound that modulates, mimics and/or blocks an olfactory
CC and/or taste sensation. GCREC DNA is useful for assessing toxicity of a
CC test compound. GCREC and its DNA are useful in the diagnosis, treatment
CC and prevention of a cell proliferative disorder e.g. actinic keratosis,
CC leukaemia, etc., a neurological disorder e.g. epilepsy, Alzheimer's
CC disease, etc., a cardiovascular disorder e.g. hypertension, angina
CC pectoris, myocardial infarction, etc., a gastrointestinal disorder e.g.
CC anorexia, cholecystitis, Crohn's disease, etc., an autoimmune/
CC inflammatory disorder e.g. acquired immune deficiency syndrome, anaemia,
CC asthma, etc., a metabolic disorder e.g. diabetes, obesity, etc., and an
CC infection by a viral agent such as adenovirus, arenavirus, etc. GCREC DNA
CC is used for creating knock out or knock in humanised animals or
CC transgenic animals to model human diseases, and somatic or germline gene
CC therapy for treating the above mentioned disorders. The present sequence
CC is human GCREC-1 protein
XX
XX Sequence 1018 AA;
SQ
Query Match 66.7%; Score 6; DB 5; Length 1018;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9
Db 213 WFLRHP 218
|||||
213 WFLRHP 218

RESULT 15
ABU07567
ID ABU07567 standard; protein; 1070 AA.
XX
XX ABU07567;
XX
XX 20-MAR-2003 (first entry)
XX
XX Human secretin type G protein-coupled receptor #1.
XX
XX Human; receptor; GPCR; G protein-coupled receptor; secretin; obesity;
XX cardiovascular disorder; diabetes; infection; HIV; pain; cancer;
XX human immunodeficiency virus infection; anorexia; bulimia; asthma;
XX Parkinson's disease; acute heart failure; hypotension; hypertension;
XX urinary retention; osteoporosis; angina pectoris; myocardial infarction;
XX ulcer; allergy; benign prostatic hypertrophy; psychosis;
XX neurological disorder; anxiety; schizophrenia; manic depression;
XX delirium; dementia; mental retardation; dyskinesia; Huntington's disease;
XX Tourette's syndrome.
XX
XX Homo sapiens.
XX
XX WO200299106-A2.
XX
XX 12-DEC-2002.
XX
XX 03-JUN-2002; 2002WO-EP006031.
XX
XX 04-JUN-2001; 2001US-0294998P.
XX
XX 26-JUL-2001; 2001US-0307608P.
XX

```

---

```

XX (FARB ) BAYER AG.
XX
XX Koehler RH, Smolyar A;
XX
XX WPI; 2003-140623/13.
DR N-PSDB; ABX15278.
XX
XX New isolated polynucleotide encoding human secretin-type G protein-
PT coupled receptor (GPCR) polypeptides, useful for preventing or treating
PT diseases associated with GPCR dysfunction, e.g. cardiovascular disease or
PT diabetes.
XX
XX Claim 1; Fig 2; 127pp; English.
XX
XX The invention relates to an isolated polynucleotide which: (a) encodes a
CC human secretin-type G protein-coupled receptor (GPCR) polypeptide; (b)
CC comprises a sequence appearing as ABX15278 and ABX15279; (c) hybridises
CC under stringent conditions to the polynucleotide in (A) and (B); (d) has
CC a sequence deviating from (A)-(C) due to the degeneration of the genetic
CC code; or represents a fragment, derivative or allelic variation of (A)-
CC (D). Also included are an expression vector containing the above
CC polynucleotide, a host cell containing the expression vector, a
CC substantially purified human secretin-type GPCR polypeptide, methods of
CC screening for agents which modulate or decrease the activity of a human
CC secretin-type GPCR, methods of reducing the activity of the human
CC secretin-type GPCR, the identified modulators. The polynucleotide is
CC useful in preventing, ameliorating, or treating diseases associated with
CC human secretin-type GPCR dysfunction. The polynucleotide may also be used
CC as hybridisation probes or primers, and in diagnostic assays or in
CC genetic testing. The methods are useful in producing and detecting the
CC polynucleotide and polypeptide and in screening for agents that modulate
CC the activity of the human secretin-type GPCR. The expression vector or
CC the reagent is useful in preparing a medicament for modulating the
CC activity of a human secretin-type GPCR in a disease, such as a
CC cardiovascular disorder, obesity, diabetes, infections (bacterial, viral,
CC fungal and protozoan), HIV (human immunodeficiency virus) infection,
CC pain, cancer, anorexia, bulimia, asthma, Parkinson's disease, acute heart
CC failure, hypotension, hypertension, urinary retention, osteoporosis,
CC angina pectoris, myocardial infarction, ulcers, allergies, benign
CC prostatic hypertrophy, psychosis, neurological disorders (e.g. anxiety,
CC schizophrenia, manic depression, delirium, dementia, mental retardation,
CC dyskinesias, Huntington's disease and Tourette's syndrome). The present
CC sequence represents a human secretin type GPCR of the invention
XX
XX Sequence 1070 AA;
SQ
Query Match 66.7%; Score 6; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9
Db 258 WFLRHP 263
|||||
258 WFLRHP 263

RESULT 16
ABG11655
ID ABG11655 standard; protein; 1131 AA.
XX
XX ABG11655;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #11646.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX

```

PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 XX DR N-PSDB; AAS75842.  
 DR  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 XX Claim 20; SEQ ID NO 42014; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activities. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AEG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 1131 AA;  
 SQ  
 Query Match 66.7%; Score 6; DB 4; Length 1131;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRHP 9  
 DB 213 WFLRHP 218  
 RESULT 17  
 ADF70474  
 ID ADF70474 standard; protein; 1232 AA.  
 XX  
 AC ADF70474;  
 XX  
 XX 12-FEB-2004 (first entry)  
 DT  
 XX Orphan receptor ligand-related human protein SeqID97.  
 DE  
 XX ligand; orphan receptor protein; fusion protein; fluorescent protein;  
 KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
 KW GFPuv; Enhanced GFP; EGFP; human.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2003071272-A1.  
 PN  
 XX 28-AUG-2003.  
 PD

XX 21-FEB-2003; 2003WO-JP001901.  
 PF  
 XX 22-FEB-2002; 2002JP-00045728.  
 PR  
 PR 23-JUL-2002; 2002JP-00213949.  
 PR 11-OCT-2002; 2002JP-00298237.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;  
 PI WPI; 2003-697654/66.  
 XX DR N-PSDB; ADF70576.  
 DR  
 XX Transformation of cells with a fusion protein of an orphan receptor  
 PT protein with a fluorescent protein useful for identification of ligands  
 PT to the orphan receptor.  
 XX  
 XX Disclosure; SEQ ID NO 97; 594pp; Japanese.  
 PS  
 XX This invention relates to a novel method of identifying ligands to an  
 CC orphan receptor protein which comprises transforming cells with DNA  
 CC encoding a fusion protein of the orphan receptor with a fluorescent  
 CC protein, so that the fusion protein is expressed in the cells (or cell  
 CC membranes isolated from them) and contacting the cells with the potential  
 CC ligand to be tested. A suitable fluorescent protein for incorporation in  
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
 CC identification of ligands binding to an orphan receptor protein.  
 XX  
 XX Sequence 1232 AA;  
 SQ  
 Query Match 66.7%; Score 6; DB 7; Length 1232;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRHP 9  
 DB 189 WFLRHP 194  
 RESULT 18  
 AEE37134  
 ID AEE37134 standard; peptide; 28 AA.  
 XX  
 AC AEE37134;  
 XX  
 XX 09-FEB-2006 (first entry)  
 DT  
 XX Human serum N-linked glycopeptide SEQ ID NO: 1238.  
 DE  
 XX Bioinformatics; blood; serum; plasma protein; protein detection;  
 KW mass spectroscopy; proteomics; glycosylation; diagnosis; cancer;  
 KW cystostatic; diabetes; antidiabetic; inflammation; antiinflammatory;  
 KW rheumatoid arthritis; antiarthritic; antirheumatic; psychiatric disorder;  
 KW neuroleptic; neurological disease; infection; antimicrobial.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2005114221-A2.  
 PN  
 XX 01-DEC-2005.  
 PD  
 XX 20-MAY-2005; 2005WO-US017842.  
 PF  
 XX 21-MAY-2004; 2004US-0573593P.  
 PR  
 XX (SYST-) INST SYSTEMS BIOLOGY.  
 PA  
 XX Aebersold RH, Zhang H;  
 PI WPI; 2006-020173/02.  
 XX  
 XX

PT Identifying glycopolypeptides in a serum or plasma sample, by identifying  
PT released sample glycopeptide fragments that correspond to standard  
PT peptides.

XX Claim 1; SEQ ID NO 1238; 193pp; English.

CC The invention relates to identifying glycopolypeptides in a serum or  
CC plasma sample comprising immobilizing derivatized sample  
CC glycopolypeptides to a solid support, releasing the sample glycopeptide  
CC fragments from the solid support, adding to the released sample  
CC glycopeptide fragments standard peptides, and identifying released sample  
CC glycopeptide fragments that correspond to standard peptides added by mass  
CC spectroscopy. Also included are a method for identifying one or more  
CC diagnostic markers for a disease, a composition comprising peptides  
CC containing the glycosylation sites (AEE35897-AEE39378), where the peptides  
CC each correspond to peptide fragments derived by cleavage of polypeptides  
CC using the same cleavage reagent) and a kit comprising peptides containing  
CC the glycosylation sites (AEE35897-AEE39378). The methods are useful for  
CC identifying glycopolypeptides in a serum or plasma sample. The methods  
CC can be used for blood serum profiling for the detection of prognostic and  
CC diagnostic protein markers. It can also be used to identify and/or  
CC validate drug targets and to evaluate drug efficacy, drug dosing, and/or  
CC drug toxicity. The methods can also be used for the detection of changes  
CC in the state of glycosylation of proteins based on the concurrent  
CC application of protein abundance measurement of protein glycosylation on  
CC the same sample. The method allows fast throughput and simplicity. It can  
CC be readily adapted for high throughput analysis of samples, which can be  
CC particularly advantageous for the analysis of clinical specimens. The  
CC method can also be automated to facilitate the processing of multiple  
CC samples. The present sequence is a human glycopeptide comprising an N-  
CC linked glycosylation site, suitable for use as a reference peptide in the  
CC method of the invention.

XX SQ Sequence 28 AA;

Query Match 55.6%; Score 5; DB 10; Length 28;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FLRHP 9  
|||||  
Db 14 FLRHP 18

RESULT 19

AAU17769  
ID AAU17769 standard; protein; 34 AA.

XX AC AAU17769;

XX DT 07-NOV-2001 (first entry)

XX DE Novel human respiratory antigen #85.

XX KW Human; respiratory antigen; respiratory disorder; throat disorder;  
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;  
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;  
KW respiratory active.

XX OS Homo sapiens.

XX PN WO200155448-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001333.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.

PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246603P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	03-DEC-2000;	2000US-0250391P.
PR	03-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX	(HUMA--)	HUMAN GENOME SCI INC.
XX	Rosen CA,	Barash SC, Ruben SM;
XX	WPI;	2001-476224/51.
DR	N-PSDB;	AAS27953.
XX	Isolated polypeptide for treating, preventing and/or prognosing	
PT	disorders related to the respiratory system including respiratory cancers	
PT	and also for testing and detection e.g. diagnosis.	
XX	Claim 11; SED ID No 387;	546pp; English.
XX	The present invention relates to the isolation of novel human respiratory	
CC	antigens, and cDNA (AAS27869-AAS28159) and genomic sequences encoding for	
CC	disorders related to the respiratory system including respiratory cancers	

CC	these polypeptides. The sequences of the invention are useful for
CC	preventing, treating and/or prognosing disorders related to the
CC	respiratory system including throat disorders (e.g. vocal cord paralysis,
CC	tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
CC	disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
CC	disorders and cancers of the respiratory tissues e.g. lung cancer. The
CC	polynucleotide sequences of the invention are useful in gene therapy and
CC	antisense therapy. AAU17685-AAU17975 represent novel human respiratory
CC	antigens. Note: The sequence data for this patent did not form part of
CC	the printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 34 AA;
	Query Match 55.6%; Score 5; DB 4; Length 34;
	Best Local Similarity 100.0%; Pred.No. 98;
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	4 WFLRH 8
D5	
	28 WFLRH 32
RESULT 20	
ADG41149	
ID	ADG41149 standard; protein; 34 AA.
XX	
AC	ADG41149;
DT	26-FEB-2004 (first entry)
DE	Human respiratory system associated protein seq id 387.
XX	
KW	antiinflammatory; antiallergic; antiasthmatic; cytostatic; gene therapy;
KW	respiratory system antigen;
KW	human respiratory system associated polynucleotide;
KW	respiratory system disorder; throat disorder; vocal cord paralysis;
KW	tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
KW	asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
KW	histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
KW	cancer; respiratory tissue cancer; throat cancer; lung cancer;
KW	cancer of the nose; gene therapy; chromosome identification; forensic;
KW	human respiratory system associated protein; human.
OS	Homo sapiens.
XX	
PW	US2003215893-A1.
XX	
PD	20-NOV-2003.
XX	
PF	07-AUG-2002; 2002US-00212872.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205515P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.
PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216880P.
PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217496P.
PR	14-JUL-2000; 2000US-0218290P.
PR	26-JUL-2000; 2000US-0220963P.
PR	26-JUL-2000; 2000US-0220964P.
PR	14-AUG-2000; 2000US-0224518P.
PR	14-AUG-2000; 2000US-0224519P.
PR	14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225575P.  
PR 14-AUG-2000; 2000US-0225578P.  
PR 14-AUG-2000; 2000US-0225579P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226661P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231142P.  
PR 08-SEP-2000; 2000US-0231143P.  
PR 08-SEP-2000; 2000US-0231144P.  
PR 08-SEP-2000; 2000US-0231143P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-02311968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234597P.  
PR 25-SEP-2000; 2000US-0234598P.  
PR 26-SEP-2000; 2000US-0234548P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 28-SEP-2000; 2000US-0235935P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241121P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-024927P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764860.  
PR 14-FEB-2002; 2002US-00074095.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-902033/82.

N-PSDB; ADG40857.

Novel respiratory system antigen and polynucleotides encoding the polypeptides, useful for treating diagnosing, treating or preventing tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat cancer.

Claim 11; SEQ ID NO 387; 236pp; English.

The invention describes an isolated polypeptide (I) comprising an amino acid sequence that is at least 90% identical to polypeptide fragment of any one of 299 respiratory system antigen sequences (PS) and having biological activity, polypeptide domain or epitope of PS, full-length protein of PS, or variant, allelic variant or species homolog of PS. (I) or a polynucleotide (II) encoding (I) is also useful for diagnosing a pathological condition or a susceptibility to a pathological condition in a subject which involves determining the presence or absence of mutation in (II) or determining the presence or amount of expression of (I) in a biological sample and diagnosing a pathological condition based on the result. The human respiratory system associated polynucleotides, the polypeptides encoded by them, and antibodies that immunospecifically bind these polypeptides are useful in diagnosis, treatment, prevention and/or prognosis of disorders of respiratory system such as throat disorders

CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders  
CC (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic  
CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,  
CC sarcooidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or  
CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and  
CC cancer of the nose). The polynucleotides are useful in gene therapy

Query Match 55.6%; Score 5; DB 7; Length 34;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
|||  
Db 28 WFLRH 32

RESULT 21  
ADI96923  
ID ADI96923 standard; peptide; 34 AA.

XX AC ADI96923;

XX 04-NOV-2004 (first entry)

XX Human respiratory system associated polypeptide SeqID387.

XX respiratory system-related polypeptide; antiasthmatic; antibacterial;  
KW antiinflammatory; cytostatic; antianaemic; antiallergic; gene therapy;  
KW pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;  
KW anaemia; leukaemia; inflammation; sinusitis;  
KW chronic obstructive pulmonary disease; infectious disease; human.

XX OS Homo sapiens.

XX US2003077704-A1.

XX 24-APR-2003.

PF 14-FEB-2002; 2002US-00074095.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 22-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764860.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM, Barash SC;  
PI WPI; 2003-765403/72.  
DR N-PSDB; ADI96631.  
XX  
XX New human respiratory system-related polypeptide and genes, useful for  
PT treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic  
PT fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or  
PT sinusitis.  
XX  
PS Claim 11; SEQ ID NO 387; 202pp; English.  
XX  
CC This invention is related to a novel isolated polypeptide, which  
CC comprises a human respiratory system-related polypeptide, and the DNA  
CC sequence which encodes it. The invention may be useful for the  
CC development of compounds with an antiasthmatic, antibacterial,  
CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In  
CC addition, the sequences disclosed may be useful for gene therapy. The  
CC polypeptide or polynucleotide is useful for treating, preventing or  
CC ameliorating a medical condition, for example pneumonia, lung cancer,  
CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,  
CC inflammations, sinusitis, chronic obstructive pulmonary disease or  
CC infectious diseases. The polypeptide or polynucleotide is also useful for  
CC diagnosing any of these diseases or a susceptibility to the disease. The  
CC present sequence is that of a human respiratory system associated  
CC polypeptide of the invention.  
XX  
XX Sequence 34 AA;  
Query Match 55.6%; Score 5; DB 7; Length 34;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 WFLRH 8  
Db 28 WFLRH 32  
|||||  
RESULT 22

ABG99963  
ID ABG99963 standard; protein; 52 AA.  
XX  
AC ABG99963;  
XX  
DT 17-JAN-2003 (first entry)  
XX  
DE Human novel polypeptide #76.  
XX  
XX Human; genetic disorder; gene mapping; medical imaging; cancer;  
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;  
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;  
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;  
KW fungal infection; bacterial infection; autoimmune disease; diabetes;  
KW atopic dermatitis.  
XX  
OS Homo sapiens.  
XX  
XX WO200274961-A1.  
PN  
XX 26-SEP-2002.  
PD  
XX 14-MAR-2002; 2002WO-US005109.  
PF  
XX 15-MAR-2001; 2001US-00810173.  
PR  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX  
DR WPI: 2003-040556/03.  
DR N-PSDB; ABX05061.  
XX  
XX New isolated polypeptides and polynucleotides, useful for preventing,  
PT treating or ameliorating medical conditions, such as cancer,  
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative  
PT disorders, and infections.  
XX  
PS Claim 9; SEQ ID NO 602; 235pp; English.  
XX  
CC The invention relates to human polynucleotides and the polypeptides they  
CC encode. The polynucleotides and polypeptides are useful in diagnostics,  
CC forensics, gene mapping, medical imaging, identification of mutations,  
CC responsible for genetic disorders or other traits, assessing biodiversity  
CC and producing many other types of data and products dependent on DNA and  
CC amino acid sequences. They are also useful for preventing, treating or  
CC ameliorating medical conditions, such as cancer, neurodegenerative  
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell  
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,  
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or  
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).  
CC Sequences ABG9988-ABG9989 and ABU0010-ABU0043 represent human  
CC polypeptides of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied by the European Patent Office  
XX  
XX Sequence 52 AA;  
Query Match 55.6%; Score 5; DB 6; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.4e-02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FLRHP 9  
Db 33 FLRHP 37  
|||||  
RESULT 23  
AAU42843  
ID AAU42843 standard; protein; 60 AA.  
XX

AC AAU42843;  
XX 27-FEB-2002 (first entry)  
XX  
XX Propionibacterium acnes immunogenic protein #3739.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
XX Propionibacterium acnes.  
OS  
XX  
XX WO200181581-A2.  
PN  
XX  
XX 01-NOV-2001.  
PD  
XX  
XX 20-APR-2001; 2001WO-US012865.  
PF  
XX  
XX 21-APR-2000; 2000US-019047P.  
PR  
XX  
XX 02-JUN-2000; 2000US-0208841P.  
PR  
XX  
XX 07-JUL-2000; 2000US-0216747P.  
PR  
XX  
XX (CORI-) CORIXA CORP.  
PA  
XX  
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
PI  
XX  
XX WPI; 2001-616774/71.  
DR  
XX  
XX N-PSDB; AAS59518.  
DR  
XX  
XX Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.  
PT  
XX  
XX Example 1; SEQ ID NO 4038; 1069pp; English.  
PS  
XX  
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 60 AA;  
Query Match 55.6%; Score 5; DB 4; Length 60;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FLRHP 9  
Db 13 FLRHP 17  
RESULT 24  
ABM39362  
XX ID ABM39362 standard; protein; 60 AA.  
XX

AC ABM39362;  
XX 20-OCT-2003 (first entry)  
XX  
XX Propionibacterium acnes predicted ORF-encoded polypeptide #4038.  
XX  
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
KW immunostimulant; immune response; vaccine.  
XX  
OS Propionibacterium acnes.  
XX  
XX WO2003033515-A1.  
PN  
XX  
XX 24-APR-2003.  
PD  
XX  
XX 11-OCT-2002; 2002WO-US032727.  
PF  
XX  
XX 15-OCT-2001; 2001US-00978825.  
PR  
XX  
XX (CORI-) CORIXA CORP.  
PA  
XX  
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
PI Barth B, Vallieue-Douglass J;  
PI  
XX  
XX WPI; 2003-381789/36.  
DR  
XX  
XX N-PSDB; ACF64447.  
DR  
XX  
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.  
XX  
XX Example 1; SEQ ID NO 4038; 1481pp; English.  
PS  
XX  
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression vectors and host cells comprising a  
CC polynucleotide of the invention; antibodies against polypeptides of the  
CC invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition comprising P. acnes polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide; a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the development of P. acnes in a  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC proteins, T cell populations or antigen-presenting cells that express the  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 60 AA;  
Query Match 55.6%; Score 5; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FLRHP 9  
Db 13 FLRHP 17



RESULT 25  
AAM86926  
ID AAM86926 standard; protein; 62 AA.  
XX AC AAM86926;  
XX DT 07-NOV-2001 (first entry)  
XX DE Human immune/haematopoietic antigen SEQ ID NO:14519.  
XX DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytotatic; gene therapy; vaccine; metastasis.  
XX OS Homo sapiens.  
XX PN WO200157182-A2.  
XX PD 09-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US0001354.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226881P.  
PR 22-AUG-2000; 2000US-0226888P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-02331968P.  
PR 14-SEP-2000; 2000US-02332397P.  
PR 14-SEP-2000; 2000US-02332398P.  
PR 14-SEP-2000; 2000US-02332399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-02334223P.  
PR 21-SEP-2000; 2000US-02334274P.  
PR 25-SEP-2000; 2000US-0233497P.  
PR 25-SEP-2000; 2000US-0233498P.  
PR 26-SEP-2000; 2000US-0233484P.  
PR 27-SEP-2000; 2000US-02335834P.  
PR 27-SEP-2000; 2000US-02335836P.  
PR 29-SEP-2000; 2000US-0233627P.  
PR 29-SEP-2000; 2000US-02336367P.  
PR 29-SEP-2000; 2000US-0233688P.  
PR 29-SEP-2000; 2000US-0233699P.  
PR 29-SEP-2000; 2000US-0233700P.  
PR 02-OCT-2000; 2000US-02336802P.  
PR 02-OCT-2000; 2000US-02337037P.  
PR 02-OCT-2000; 2000US-02337038P.  
PR 02-OCT-2000; 2000US-02337039P.  
PR 02-OCT-2000; 2000US-02337040P.  
PR 13-OCT-2000; 2000US-0233935P.  
PR 13-OCT-2000; 2000US-0233937P.  
PR 20-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.

PR	05-DEC-2000;	2000US-0251030P.	
PR	05-DEC-2000;	2000US-0251988P.	
PR	06-DEC-2000;	2000US-0256719P.	
PR	06-DEC-2000;	2000US-0251479P.	
PR	08-DEC-2000;	2000US-0251856P.	
PR	08-DEC-2000;	2000US-0251868P.	
PR	08-DEC-2000;	2000US-0251869P.	
PR	08-DEC-2000;	2000US-0251989P.	
PR	08-DEC-2000;	2000US-0251990P.	
PR	11-DEC-2000;	2000US-0254097P.	
PR	05-JAN-2001;	2001US-0259678P.	
XX	(HUMA-)	HUMAN GENOME SCI INC.	
XX	Rosen CA,	Barash SC, Ruben SM;	
XX	WPI;	2001-483426/52.	
DR	N-PSDB;	AAK59707.	
XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,		
PT	useful for preventing, diagnosing and/or treating cancers and metastasis.		
XX	Claim 11; SEQ ID NO 14519;	3071pp + Sequence Listing; English.	
XX	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)		
CC	amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic		
CC	activity, and can be used in gene therapy and vaccine production. (I)		
CC	proteins and polynucleotides may be used in the prevention, diagnosis and		
CC	treatment of diseases associated with inappropriate (I) expression. For		
CC	example, they may be used to treat disorders associated with decreased		
CC	expression by rectifying mutations or deletions in a patient's genome		
CC	that affect the activity of (I) by expressing inactive proteins or to		
CC	supplement the patients own production of (I). Additionally, (I)		
CC	polynucleotides may be used to produce the secreted (I), by inserting the		
CC	nucleic acids into a host cell and culturing the cell to express the		
CC	protein. (I) proteins and polynucleotides may be used to prevent,		
CC	diagnose and treat immune/hematopoietic-related diseases, especially		
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703		
CC	to AAK87694 represent human immune/hematopoietic antigen genomic		
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK62169		
CC	represent sequences used in the exemplification of the present invention		
XX	Sequence 62 AA;		
SQ	Query Match	55.6%; Score 5; DB 4; Length 62;	
	Best Local Similarity	100.0%; Pred. No. 1.7e-02;	
	Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	2 ETWFL 6		
Db			
	39 ETWFL 43		
RESULT 26			
AAQ98737			
ID	AAQ98737 standard; protein; 64 AA.		
XX	AAQ98737;		
AC	AAQ98737;		
XX	21-SEP-2001 (first entry)		
DT	Human cell death protective cDNA clone CNI-00720 ORF3 protein, SEQ:267.		
XX	Cell death protective; apoptosis; necrosis; human; drug screening;		
KW	cell death-associated disorder; central nervous system disorder;		
KW	psychiatric disorder; neurological disorder; ischaemia-related disorder;		
KW	stroke; cerebral infarction; ischaemic encephalopathy;		
KW	neurodegenerative disorder; Alzheimer's disease; Huntington's disease;		
KW	Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;		
KW	vascular disease; ophthalmological disorder; diabetic retinopathy;		
KW	macular degeneration; hypertension; myocardial infarction;		
KW	atherosclerosis; respiratory disorder; asthma; transgenic animal;		
KW	chronic obstructive pulmonary disease; neoplastic condition; cancer;		
KW	benign tumour; anaemia; gastrointestinal disorder; gastritis;		
KW	ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;		
KW	glomerulonephritis; cystitis; endometritis; endocrine disorder;		
KW	Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;		
KW	urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.		
OS	Homo sapiens.		
XX	WO200145638-A2.		
FN	28-JUN-2001.		
XX	11-DEC-2000;	2000WO-US033547.	
PF	14-DEC-1999;	99US-00461697.	
XX	(COGE-)	COGENT NEUROSCIENCE INC.	
PA	Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;		
PI	WPI;	2001-390297/41.	
XX	N-PSDB;	AAH84265, AAH84268.	
DR	Novel protective sequence polynucleotides and polypeptides, used to		
PT	identify modulators of their expression and activity, which are used in		
PT	to treat central nervous system conditions, diseases and disorders.		
XX	Claim 1; Fig 10C; 325pp; English.		
XX	Sequences AAH84132-AAH84370 represent human nucleic acid sequences which		
CC	protect against cell death (i.e., apoptosis or necrosis). Sequences		
CC	AAH84132, AAH84145, AAH84170, AAH84201, AAH84226, AAH84265,		
CC	AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,		
CC	while the remaining nucleic acid sequences within the range given above		
CC	represent the open reading frames (ORFs) of these cDNA clones. Sequences		
CC	AAQ98610-AAQ98829 represent the polypeptides encoded by the cell death		
CC	protective ORFs. The cell death protective cDNA clones are able to		
CC	prevent, delay or reverse progression through the apoptotic or necrotic		
CC	pathways when injected into a cell predisposed to or undergoing cell		
CC	death. The cell death protective nucleic acids and polypeptides can be		
CC	used in the diagnosis and treatment of disorders associated with cell		
CC	death, and to screen for compounds which modulate their activity or		
CC	expression. Such modulators, preferably a small organic molecule, an		
CC	antibody, a ribozyme, or an antisense molecule, can also be used to treat		
CC	cell death-related diseases. Such diseases include those associated with		
CC	the central nervous system including psychiatric or neurological		
CC	disorders, especially ischaemia-related conditions such as strokes, and		
CC	also includes neurodegenerative disorders such as Alzheimer's disease,		
CC	Huntington's disease, or Parkinson's disease. The modulators may also be		
CC	used to treat infections such as meningitis, malaria, or trypanosomiasis;		
CC	vascular diseases such as ischaemic encephalopathy or cerebral infarction;		
CC	; eye conditions such as diabetic retinopathy or macular degeneration;		
CC	hypertension; myocardial infarction; atherosclerosis; respiratory		
CC	conditions such as asthma or chronic obstructive pulmonary disease;		
CC	neoplastic conditions such as cancers or benign tumours; blood cell		
CC	conditions such as anaemia; gastrointestinal conditions such as gastritis		
CC	or ulcerative colitis; liver conditions such as biliary cirrhosis; endocrine		
CC	disorders such as glomerulonephritis; cystitis; endometritis; skin		
CC	disorders such as Grave's disease or Hashimoto's thyroiditis; AIDS		
CC	conditions such as dermatitis or urticaria; or immune system disorders		
CC	such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may		
CC	additionally be used to generate animal models of cell death-associated		
CC	disorders. The present sequence represents a cell death protective		
CC	polypeptide		
XX	Sequence 64 AA;		
SQ	Query Match	55.6%; Score 5; DB 4; Length 64;	
	Best Local Similarity	100.0%; Pred. No. 1.7e+02;	
	Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	4 WFLRH 8		

Db	24	WFLRH	28
RESULT 27			
AAU50032			
ID	AAU50032	standard; protein; 64 AA.	
XX	AAU50032;		
XX	27-FEB-2002	(first entry)	
XX	Propionibacterium	acnes immunogenic protein #10928.	
XX	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;		
KW	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;		
KW	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;		
KW	dermatological; osteopathic; neuroprotectant.		
XX	Propionibacterium	acnes.	
OS	WO200181581-A2.		
XX	01-NOV-2001.		
XX	20-APR-2001; 2001WO-US012865.		
XX	21-APR-2000; 2000US-0199047P.		
PR	02-JUN-2000; 2000US-0208841P.		
PR	07-JUL-2000; 2000US-0216747P.		
XX	(CORI-) CORIXA CORP.		
PA	Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;		
XX	L'maisonneuve J, Zhang Y, Jen S, Carter D;		
PI	WFI; 2001-616774/71.		
DR	N-PSDB; AAS95946.		
DR	Propionibacterium	acnes polypeptides and nucleic acids useful for	
XX	vaccinating against and diagnosing infections, especially useful for		
PT	treating acne vulgaris.		
PT	Example 1; SEQ ID NO 11227; 1069pp; English.		
XX	Sequences AAU39105-AAU68017 represent Propionibacterium	acnes immunogenic	
CC	polypeptides. The proteins and their associated DNA sequences are used in		
CC	the treatment, prevention and diagnosis of medical conditions caused by		
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,		
CC	pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.		
CC	P. acnes is also involved in infections of bone, joints and the central		
CC	nervous system, however it is particularly involved in the inflammatory		
CC	lesions associated with acne vulgaris. A method for detecting the		
CC	presence or absence of P. acnes in a patient comprises contacting a		
CC	sample with a binding agent that binds to the proteins of the invention		
CC	and determining the amount of bound protein in the sample. The		
CC	polypeptides may be used as antigens in the production of antibodies		
CC	specific for P. acnes proteins. These antibodies can be used to		
CC	downregulate expression and activity of P. acnes polypeptides and		
CC	therefore treat P. acnes infections. The antibodies may also be used as		
CC	diagnostic agents for determining P. acnes presence, for example, by		
CC	enzyme linked immunosorbent assay (ELISA). Note: The sequence data for		
CC	this patent did not form part of the printed specification, but was		
CC	obtained in electronic format directly from WIPO at		
CC	ftp.wipo.int/pub/published_pct_sequences		
XX	Sequence 64 AA;		
SQL	Query Match	55.6%;	Score 5; DB 4; Length 64;
	Best Local Similarity	100.0%;	Pred.No. 1.7e+02;
	Matches	5; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	5	FLRHP	9

Db	30 FLRHP 34
RESULT 28	
ABM46551	
ID	ABM46551 standard; protein; 64 AA.
XX	AC
XX	ABM46551;
XX	
DT	20-OCT-2003 (first entry)
XX	
XX	Propionibacterium acnes predicted ORF-encoded polypeptide #11227.
DE	
XX	Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW	immunostimulant; immune response; vaccine.
XX	
OS	Propionibacterium acnes.
XX	
PN	WO2003033515-A1.
XX	
PD	24-APR-2003.
XX	
PF	11-OCT-2002; 2002WO-US032727.
XX	
PR	15-OCT-2001; 2001US-00978825.
XX	(CORI-) CORIXA CORP.
PA	
PI	Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI	Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI	Barth B, Valliee-Douglas J;
XX	
DR	WPI; 2003-381789/36.
DR	N-PSDB; ACF64475.
XX	
PT	New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT	polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT	or for stimulating an immune response specific for a P. acnes protein.
XX	
PS	Example 1; SEQ ID NO 11227; 1481pp; English.
XX	
CC	The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC	encoding a Propionibacterium acnes protein. The invention also relates to
CC	polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC	immunogenic fragments of P. acnes polypeptides. The invention
CC	additionally encompasses expression vectors and host cells comprising a
CC	polynucleotide of the invention; antibodies against polypeptides of the
CC	invention; fusion proteins comprising a polypeptide of the invention; a
CC	method for stimulating an immune response specific for a P. acnes
CC	polypeptide and an isolated T cell population comprising T cells prepared
CC	via this method; a vaccine composition (comprising P. acnes polypeptides,
CC	polynucleotides, antibodies, fusion proteins, T cell populations, or
CC	antigen-presenting cells that express the polypeptide); a method and kit
CC	for detecting or determining the presence or absence of P. acnes in a
CC	patient; and a method for inhibiting the development of P. acnes in a
CC	patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC	proteins, T cell populations or antigen-presenting cells that express the
CC	polypeptides are useful for diagnosing, preventing or treating acne
CC	vulgaris, or for stimulating an immune response specific for a P. acnes
CC	protein. The polynucleotides can also be used as probes or primers for
CC	nucleic acid hybridisation. The vaccine composition is useful for the
CC	stimulation of an immune response against P. acnes, or for treating acne,
CC	and the kit is useful for performing a diagnostic assay. The present
CC	sequence represents a polypeptide predicted to be encoded by an ORF (open
CC	reading frame) contained within the P. acnes polynucleotides of the
CC	invention. Note: The sequence data for this patent did not form part of
CC	the printed specification, but was obtained in electronic format directly
CC	from WIPO at <a href="http://fp.wipo.int/pub/published_pct_sequences">fp.wipo.int/pub/published_pct_sequences</a>
XX	
SQ	Sequence 64 AA;
Query Match	55.6%; Score 5; DB 6; Length 64;
Best Local Similarity	100.0%; Pred. No. 1.7e+02;

Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	5 FLRHP 9								
Db	30 FLRHP 34								
RESULT 29									
ABP03674									
ID	ABP03674	standard; protein; 73 AA.							
XX	AC	ABP03674;							
XX	DT	25-JUN-2002 (first entry)							
XX	DE	Human ORFX protein sequence SEQ ID NO:7330.							
XX	KW	Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis; hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; degenerative disorder; osteoarthritis; neurodegenerative disorder; cardiovascular disease; diabetes mellitus; systemic lupus erythematosus; hypotension; hypothyroidism; cholesterol ester storage disease; immune deficiency; immune disorder; infectious disease; autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis; myasthenia gravis.							
XX	OS	Homo sapiens.							
XX	PN	WO200192523-A2.							
XX	PD	06-DEC-2001.							
XX	XX	29-MAY-2001; 2001WO-US010836.							
XX	PR	30-MAY-2000; 2000US-0206132P.							
XX	PR	29-AUG-2000; 2000US-0228716P.							
XX	PA	(CURA-) CURAGEN CORP.							
XX	PI	Shinkets RA, Leach MD;							
XX	DR	WPI; 2002-106308/14.							
XX	DR	N-PSDB; ABN19426.							
XX	PT	Novel human polypeptides and polynucleotides useful for diagnosing, preventing and treating cardiovascular disease, neurodegenerative, hyperproliferative disorders and autoimmune disorders.							
XX	PS	Disclosure; SEQ ID NO 7330; 1037pp; English.							
XX	CC	The present invention describes substantially purified human proteins (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1 in the specification). ABN15762 to ABN27252 encode the human ORFX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for treating or preventing a pathology associated with an ORFX-associated disorder in humans, and in the manufacture of a medicament for treating a syndrome associated with ORFX-associated disorder. ORFX polynucleotide sequences can be used in gene therapy. ORFX sequences can be used in the treatment of cancer, hyperproliferative disorders, cirrhosis of liver, psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorders, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORFX proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues and conditions resulting from systemic cytokine damage. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic							

CC	Format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences									
XX	Sequence 73 AA;									
SQ	Query Match	55.6%;	Score 5;	DB 5;	Length 73;					
	Best Local Similarity	100.0%;	Pred. No. 1.9e+02;							
	Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	5 FLRHP 9									
Db	3 FLRHP 7									
RESULT 30										
AAG98736										
ID	AAG98736	standard; protein; 75 AA.								
XX	AC	AAG98736;								
XX	DT	21-SEP-2001 (first entry)								
XX	DE	Human cell death protective cDNA clone CNI-00720	ORF2 protein, SEQ:265.							
XX	KW	Cell death protective; apoptosis; necrosis; human; drug screening; cell death-associated disorder; central nervous system disorder; psychiatric disorder; neurological disorder; ischaemia-related disorder; stroke; cerebral infarction; ischaemic encephalopathy; neurodegenerative disorder; Alzheimer's disease; Huntington's disease; Parkinson's disease; infection; meningitis; malaria; trypanosomiasis; vascular disease; ophthalmological disorder; diabetic retinopathy; macular degeneration; hypertension; myocardial infarction; atherosclerosis; respiratory disorder; asthma; transgenic animal; chronic obstructive pulmonary disease; neoplastic condition; cancer; benign tumour; anaemia; gastrointestinal disorder; gastritis; ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder; glomerulonephritis; cystitis; endometriosis; endocrine disorder; Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis; urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.								
OS	Homo sapiens.									
XX	PN	WO200145638-A2.								
XX	PD	28-JUN-2001.								
XX	XX	11-DEC-2000; 2000WO-US033547.								
XX	PR	14-DEC-1999; 99US-00461697.								
XX	PA	(COGE-) COGENT NEUROSCIENCE INC.								
XX	PI	Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;								
XX	DR	WPI; 2001-390297/41.								
XX	DR	N-PSDB; AAH84265, AAH84267.								
XX	PT	Novel protective sequence polynucleotides and polypeptides, used to identify modulators of their expression and activity, which are used in to treat central nervous system conditions, diseases and disorders.								
XX	PS	Claim 1; Fig 10B; 325pp; English.								
XX	CC	Sequences AAH84132-AAH84370 represent human nucleic acid sequences which protect against cell death (i.e., apoptosis or necrosis). Sequences AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265, AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones, while the remaining nucleic acid sequences within the range given above represent the open reading frames (ORFs) of these cDNA clones. Sequences AAG98610-AAG98829 represent the polypeptides encoded by the cell death protective ORFs. The cell death protective cDNA clones are able to prevent, delay or reverse progression through the apoptotic or necrotic pathways when injected into a cell predisposed to or undergoing cell death. The cell death protective nucleic acids and polypeptides can be								

CC used in the diagnosis and treatment of disorders associated with cell  
 CC death, and to screen for compounds which modulate their activity or  
 CC expression. Such modulators, preferably a small organic molecule, an  
 CC antibody, a ribozyme, or an antisense molecule, can also be used to treat  
 CC cell death-related diseases. Such diseases include those associated with  
 CC the central nervous system including psychiatric or neurological  
 CC disorders, especially ischaemia-related conditions such as strokes, and  
 CC also includes neurodegenerative disorders such as Alzheimer's disease,  
 CC Huntington's disease, or Parkinson's disease. The modulators may also be  
 CC used to treat infections such as meningitis, malaria, or trypanosomiasis;  
 CC vascular diseases such as ischaemic encephalopathy or cerebral infarction  
 CC; eye conditions such as diabetic retinopathy or macular degeneration;  
 CC hypertension; myocardial infarction; atherosclerosis; respiratory  
 CC conditions such as asthma or chronic obstructive pulmonary disease;  
 CC neoplastic conditions such as cancers or benign tumours; blood cell  
 CC conditions such as anaemia; gastrointestinal conditions such as gastritis  
 CC or ulcerative colitis; liver conditions such as biliary cirrhosis, kidney  
 CC disorders such as glomerulonephritis; cystitis; endometriosis; skin  
 CC disorders such as Grave's disease or Hashimoto's thyroiditis; skin  
 CC conditions such as dermatitis or urticaria; or immune system disorders  
 CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may  
 CC additionally be used to generate animal models of cell death-associated  
 CC disorders. The present sequence represents a cell death protective  
 CC polypeptide

XX  
 SQ Sequence 75 AA;

Query Match 55.6%; Score 5; DB 4; Length 75;

Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 0;

QY 4 WFLRH 8  
 |||||  
 Db 35 WFLRH 39

RESULT 31

AA999844  
 ID AA999844 standard; protein; 76 AA.

AC AA999844;

DT 07-JAN-2002 (first entry)

DE Human excretory related polypeptide SEQ ID NO 581.

DE Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
 KW antiparkinsonian; antiskickling; antianaemic; antiarthritic; cancer;  
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;  
 KW antiparasitic; cardiatic; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine;  
 KW excretory system.

XX Homo sapiens.

XX WO200155313-A2.

XX 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US001323.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180828P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.



PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225577P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 22-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241122P.  
PR 20-OCT-2000; 2000US-0241185P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244647P.  
PR 08-NOV-2000; 2000US-0244647P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI: 2001-488784/53.  
XX N-PSDB; AA163213.  
XX  
XX New isolated nucleic acids and polypeptides, useful for diagnosing, treating and/or preventing human diseases and disorders.  
PT  
XX  
XX Claim 11; SEQ ID NO 528; 564pp + Sequence Listing; English.  
XX  
XX The invention relates to novel kidney related polynucleotides (AA162971-AA163793) and the encoded polypeptides (AA42417-AA42691) collectively known as kidney antigens and the use of such kidney antigens for detecting disorders of the kidney, especially kidney cancer and kidney cancer metastases. The polynucleotides and proteins are also useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at



```

CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 76 AA;

Query Match      55.6%; Score 5; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7
DB 4 TWFLR 8
|||||

RESULT 33
ABP34862
ID ABP34862 standard; protein; 82 AA.
AC
XX ABP34862;
DT 08-JUL-2002 (first entry)
XX
DE Human ORF3835 protein, SEQ ID NO:7670.
XX
KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neutropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS Homo sapiens.
XX
XX WO200190366-A2.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-US017076.
XX
XX 24-MAY-2000; 2000US-0206690P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Leach MD, Shinkets RA;
XX
XX WPI; 2002-106200/14.
XX
XX N-PSDB; ABN78898.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and disorders related to organ
XX transplantation.
XX
XX Claim 10; Page 2162; 2508pp; English.
XX
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
XX designated ORF (open reading frame) 1-4534, and sequences ABN75054-
XX ABN79587 represent cDNAs encoding them. The invention also encompasses
XX polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
XX referred to as ORFX) proteins, polynucleotides at least 85% identical to
XX the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
XX polynucleotides, the recombinant production of ORFX proteins, antibodies
XX specific for ORFX proteins, methods of detecting ORFX polynucleotides and
XX polypeptides, methods of screening for modulators of ORFX expression or
XX activity, and methods of screening individuals for a predisposition to an
XX ORFX-associated disorder. The ORFX proteins of the invention have a wide
XX range of biological activities, such as cytokine, cell proliferation,

```

---

```

CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antifective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases
XX
XX Sequence 82 AA;
SQ
Query Match      55.6%; Score 5; DB 5; Length 82;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7
DB 76 TWFLR 80
|||||

RESULT 34
ADT58131
ID ADT58131 standard; protein; 85 AA.
XX
XX AC ADT58131;
XX
XX 13-JAN-2005 (first entry)
XX
XX Plant polypeptide, SEQ ID 8208.
XX
XX DE
XX
XX KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;
XX disease resistance; galactomannan production; plant growth regulator;
XX heat tolerance; herbicide tolerance; lignin production;
XX extreme osmotic condition tolerance; pathogens resistance;
XX pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
XX OS Viridiplantae.
XX
XX XX US2004216190-A1.
XX
XX 28-OCT-2004.
XX
XX PF 18-DEC-2003; 2003US-00739930.
XX
XX PR 28-APR-2003; 2003US-00424599.
XX
XX PR 28-APR-2003; 2003US-00425115.
XX
XX (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
XX
XX WPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
XX genetics, and in particular for producing transgenic plants with improved
XX biological characteristics.
XX
XX Claim 2; SEQ ID NO 8208; 14pp; English.
XX

```



CC The invention relates a recombinant DNA construct comprising a  
 CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:  
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences  
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,  
 CC Arabidopsis, wheat and rape but the specification does not indicate which  
 CC sequences is derived from which organism. Also included is a method of  
 CC producing a plant having an improved property, comprising transforming a  
 CC plant with a recombinant DNA construct comprising a promoter region  
 CC functional in a plant cell operably joined to a polynucleotide encoding a  
 CC polypeptide associated with the property, and growing the transformed  
 CC plant. The property is selected from improving plant cold tolerance, for  
 CC manipulating growth rate in plant cells by modification of the cell cycle  
 CC pathway, for improving plant drought tolerance, for providing increased  
 CC resistance to plant disease, for galactomannan production, for production  
 CC of plant growth regulators, for improving plant heat tolerance, for  
 CC improving plant tolerance to herbicides, for increasing the rate of  
 CC homologous recombination in plants, for lignin production, for improving  
 CC plant tolerance to extreme osmotic conditions, for improving plant  
 CC tolerance to pathogens or pests, for yield improvement by modification of  
 CC photosynthesis, for modifying seed oil yield and/or content, for  
 CC modifying seed protein yield and/or content, for yield improvement by  
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC and for yield improvement by providing improved plant growth and  
 CC development under at least one stress condition. The polynucleotide may  
 CC also encode a plant transcription factor. The methods and compositions of  
 CC the present invention are useful in the field of biochemistry and  
 CC genetics, in particular for producing transgenic plants with improved  
 CC biological characteristics such as increased yield, improved nitrogen  
 CC flow, increasing plant tolerance to cold or heat, improving plant  
 CC tolerance to extreme osmotic and drought conditions, and improving plant  
 CC tolerance to plant pests or pathogens. They can also be used in physical  
 CC arrays of molecules, plant breeding markers, computer-based storage and  
 CC analysis systems. The present sequence is one of the 5544 plant protein  
 CC sequences of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=20040216190.  
 XX  
 SQ Sequence 85 AA;

Query Match 55.6%; Score 5; DB 8; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 |||||  
 Db 14 FLRHP 18

RESULT 35  
 AAU51518  
 ID AAU51518 standard; protein; 86 AA.  
 XX  
 AC AAU51518;

XX  
 DT 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #12414.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US012865.

XX PR 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.  
 PR 07-JUL-2000; 2000US-0216747P.  
 XX  
 PA (CORI-) CORIXA CORP.

XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX WPI; 2001-616774/71.  
 DR N-PSDB; AAS59551.

XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris.

PS Example 1; SEQ ID NO 12713; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 86 AA;

Query Match 55.6%; Score 5; DB 4; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 |||||  
 Db 2 FLRHP 6

RESULT 36  
 ABM48037  
 ID ABM48037 standard; protein; 86 AA.  
 XX  
 AC ABM48037;

XX  
 DT 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #12713.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
 KW immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX PD 24-APR-2003.

XX PF 11-OCT-2002; 2002WO-US032727.

XX PR 15-OCT-2001; 2001US-00978825.

XX PA (CORI-) CORIXA CORP.

XX Mitcam JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
 PI Zhang Y, Wang S, Jen S, Lodes WJ, Benson DR, Jones R, Carter D;  
 PI Barth B, Vallieue-Douglas J;  
 XX  
 DR WPI: 2003-381789/36.  
 DR N-PSDB; ACF64480.  
 XX  
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a P. acnes protein.  
 XX  
 XX Example 1; SEQ ID NO 12713; 1481pp; English.  
 XX  
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
 CC encoding a Propionibacterium acnes protein. The invention also relates to  
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
 CC immunogenic fragments of P. acnes polypeptides. The invention  
 CC additionally encompasses expression vectors and host cells comprising a  
 CC polynucleotide of the invention; antibodies against polypeptides of the  
 CC invention; fusion proteins comprising a polypeptide of the invention; a  
 CC method for stimulating an immune response specific for a P. acnes  
 CC polypeptide and an isolated T cell population comprising T cells prepared  
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
 CC antigen-presenting cells that express the polypeptide); a method and kit  
 CC for detecting or determining the presence or absence of P. acnes in a  
 CC patient; and a method for inhibiting the development of P. acnes in a  
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
 CC proteins, T cell populations or antigen-presenting cells that express the  
 CC polypeptides are useful for diagnosing, preventing or treating acne  
 CC vulgaris, or for stimulating an immune response specific for a P. acnes  
 CC protein. The polynucleotides can also be used as probes or primers for  
 CC nucleic acid hybridisation. The vaccine composition is useful for the  
 CC stimulation of an immune response against P. acnes, or for treating acne,  
 CC and the kit is useful for performing a diagnostic assay. The present  
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
 CC reading frame) contained within the P. acnes polynucleotides of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)  
 XX  
 XX Sequence 86 AA;  
 SQ  
 Query Match 55.6%; Score 5; DB 6; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRHP 9  
 Db 2 FLRHP 6  
 |||||  
 RESULT 37  
 ADX94950  
 ID ADX94950 standard; protein; 89 AA.  
 XX  
 AC ADX94950;  
 XX  
 DT 21-APR-2005 (first entry)  
 XX  
 DE Plant full length insert polypeptide seqid 57614.  
 XX  
 KW plant protectant; plant growth regulant; gene therapy; plant;  
 KW recombinant DNA construct; physical array; plant breeding marker;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;  
 KW growth rate; cell cycle pathway; disease resistance;  
 KW Galactomannan production; lignin production; plant growth regulator;  
 KW yield; plant growth; plant development; seed oil; protein yield;  
 KW protein content.  
 XX  
 OS Unidentified.

XX US2004034888-A1.  
 FN  
 XX 19-FEB-2004.  
 PD  
 XX  
 XX 28-APR-2003; 2003US-00425114.  
 PF  
 XX  
 PR 06-MAY-1999; 99US-00304517.  
 PR 05-NOV-2001; 2001US-00985678.  
 XX  
 PA (LIUJ/) LIU J.  
 PA (ZHOU/) ZHOU Y.  
 PA (KOVA/) KOVALIC D K.  
 PA (SCRE/) SCREEN S E.  
 PA (TABA/) TABASKA J E.  
 PA (CAOY/) CAO Y.  
 XX  
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;  
 DR WPI: 2004-180133/17.  
 XX  
 XX New recombinant DNA construct, useful for improving plant tolerance to  
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
 PT pests, for conferring increased resistance to plant disease, or for  
 PT improving yield.  
 XX  
 PS Claim 1; SEQ ID NO 57614; 15pp; English.  
 XX  
 XX The invention describes a recombinant DNA construct comprising a  
 CC polynucleotide consisting of a sequence encoding an amino acid sequence  
 CC available in electronic form from the US patent office at  
 CC [ftp.secdat.uspto.gov/sequence.html?DocID:2004034888](http://ftp.secdat.uspto.gov/sequence.html?DocID:2004034888). The polynucleotide  
 CC of the invention are also useful in physical arrays of molecules and as  
 CC plant breeding markers. The recombinant DNA construct is useful for  
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, for conferring  
 CC increased resistance to plant disease, for producing galactomannan,  
 CC lignin or plant growth regulators, for increasing the rate of homologous  
 CC recombination in plants, for improving yield by modification of  
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC or by providing improved plant growth and development under at least one  
 CC stress condition or for modifying seed oil or protein yield and/or  
 CC content. This is the amino acid sequence of a plant full length insert  
 CC polypeptide that can be used in the recombinant DNA construct of the  
 CC invention.  
 XX  
 SQ Sequence 89 AA;  
 Query Match 55.6%; Score 5; DB 8; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRH 8  
 Db 6 WFLRH 10  
 |||||  
 RESULT 38  
 ADK36992  
 ID ADK36992 standard; protein; 90 AA.  
 XX  
 AC ADK36992;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Novel human polypeptide SeqID9074.  
 XX  
 KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;  
 KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;  
 KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;  
 KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;  
 KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;

KW	fungus; parasite; human.
XX	
OS	Homo sapiens.
XX	
PH	Key Location/Qualifiers
FT	Misc-difference 1: .90 /label= OTHER
FT	/notes =OTHER= All Xaa's in this sequence are unknown amino acids or the site of a stop codon within the DNA sequence"
FT	
FT	
FT	
FT	
XX	WO200216439-A2.
PN	
XX	
PD	28-FEB-2002.
XX	
XX	
XX	05-MAR-2001; 2001WO-US004941.
XX	
PR	07-MAR-2000; 2000US-00519705.
PR	19-MAY-2000; 2000US-00574454.
XX	(HYSE-) HYSEQ INC.
PA	
PI	Tang YT, Liu C, Drmanac RT;
XX	
DR	WPI; 2002-280918/32.
XX	
PT	Isolated polynucleotide encoding bone marrow derived polypeptides useful for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's disease, and inflammatory bowel disease.
PT	
PT	
PT	
XX	Claim 20; SEQ ID NO 9074; 504pp; English.
PS	
XX	
CC	This invention relates to a novel isolated polynucleotide comprising a nucleotide sequence selected from one of 1680 sequences, a mature protein coding portion of them, an active domain of them and their complementary sequences. The invention may be useful for the production of compounds with an antiarthritic, antiparkinsonian, neuroprotective, nontropic, immunosuppressive, cytostatic, antipsoriatic, antiinflammatory, antibacterial, antiviral, antifungal or antiparasitic activity. In addition, the disclosed sequences may be useful for gene therapy. The polypeptides or their antibodies are useful for treating many diseases such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer, psoriasis, inflammatory bowel disease and infections caused by bacteria, viruses, fungi or parasites. The present sequence is that of a human polypeptide of the invention.
CC	
XX	Sequence 90 AA;
SQ	
Query Match	55.6%; Score 5; DB 5; Length 90;
Best Local Similarity	100.0%; Pred. No. 2.3e+02;
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 VETWF 5 
Db	58 VETWF 62
RESULT 39	
ABM94143	ID ABM94143 standard; protein; 98 AA.
XX	
AC	ABM94143;
XX	
DT	02-JUN-2005 (first entry)
XX	
DE	M. xanthus protein sequence, seq id 13342.
XX	
KW	Transgenic plant; DNA replication; gene regulation; gene expression.
XX	
OS	Myxococcus xanthus.
XX	
PN	US6833447-B1.
XX	
PD	21-DEC-2004.
XX	
Pf	10-JUL-2001; 2001US-00902540.
XX	
PR	10-JUL-2000; 2000US-0217883P.
XX	(MONS ) MONSANTO TECHNOLOGY LLC.
PA	
XX	Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
PI	
XX	WPI; 2005-028716/03.
DR	
XX	
PT	New substantially purified Myxococcus xanthus nucleic acid molecule encoding a nitrite reductase, useful for determining gene expression, identifying mutations in a gene of interest, and for constructing mutations in a gene of interest.
PT	
PT	
XX	Example 2; SEQ ID NO. 13342; 25pp; English.
PS	
XX	
CC	The invention relates to a substantially purified nucleic acid molecule encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a recombinant DNA construct for expression of a nitrite reductase gene in a plant cell, and a plant cell comprising the recombinant DNA construct. The nucleic acid is useful for determining gene expression, identifying mutations in a gene of interest, and for constructing mutations in a gene of interest. Sequences given in records for SEQ IDs 9692-16825 represent a group of 7134 Myxococcus xanthus proteins and peptides. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO
CC	
XX	Sequence 98 AA;
SQ	
Query Match	55.6%; Score 5; DB 9; Length 98;
Best Local Similarity	100.0%; Pred. No. 2.5e+02;
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	5 FLRHP 9 
Db	29 FLRHP 33
RESULT 40	
ADC14235	ID ADC14235 standard; protein; 102 AA.
XX	
AC	ADC14235;
XX	
DT	18-DEC-2003 (first entry)
XX	
DE	Human enzyme ENZM-41.
XX	
KW	enzyme; human; ENZM; cytosolic; antiarteriosclerotic; antidiabetic; anticongestant; nontropic; neuroprotective; cerebroprotective; anti-HIV; antiallergic; antiinflammatory; thyromimetic; gene therapy;
KW	cell proliferative disorder; endocrine disorder; neurological disorder; immune system disorder; inflammatory disorder; developmental disorder; reproductive disorder; vesicle-trafficking disorder; infection.
XX	
OS	Homo sapiens.
XX	
PN	WO2003042357-A2.
XX	
PD	22-MAY-2003.
XX	
Pf	26-SEP-2002; 2002WO-US031096.
XX	
PR	28-SEP-2001; 2001US-0326388P.
PR	12-OCT-2001; 2001US-0328979P.
PR	19-OCT-2001; 2001US-0346034P.
PR	26-OCT-2001; 2001US-0348284P.
PR	08-NOV-2001; 2001US-0338048P.
PR	16-NOV-2001; 2001US-0332340P.
PR	14-DEC-2001; 2001US-0340357P.

PR 29-MAR-2002; 2002US-0368722P.  
PR 29-MAR-2002; 2002US-0368799P.  
PR 17-MAY-2002; 2002US-0381558P.  
PR 07-JUN-2002; 2002US-0387119P.  
PR 21-JUN-2002; 2002US-0390662P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Yang J, Lu DAM, Yue H, Elliott VS, Warren BA, Dugan BM;  
PI Forsythe IJ, Lee EA, Hafalia AJA, Ramkumar J, Chawla NK, Baughn MR;  
PI Becha SD, Gorvad AE, Tran UK, Li JX, Yao MG, Ison CH, Griffin JA;  
PI Lee SY, Chang H, Emerling BM, Tang YT, Lal PG, Kable AE;  
PI Marquis JP, Jiang X, Jackson AA, Zebajadian Y, Swarnakar A;  
PI Wilson AD, Jin P, Richardson TW, Bhatia U, Burrill JD, Lee S;  
PI Blake JJ, Ho A, Zheng W, Gao J;  
XX  
XX WPI; 2003-449567/42.  
DR N-PSDB; ADC14288.  
XX  
XX New human enzymes (ENZM), useful for diagnosing, treating and preventing  
PT diseases or conditions associated with the aberrant ENZM expression e.g.  
PT cancer, diabetes, epilepsy, or infections.  
XX  
XX Claim 1; SEQ ID NO 41; 416pp; English.  
XX  
XX The invention relates to a novel isolated human enzyme (ENZM)  
CC polypeptide. A polypeptide of the invention has cytostatic,  
CC antiarteriosclerotic, antidiabetic, anticonvulsant, nootropic,  
CC neuroprotective, cerebroprotective, anti-HIV, antiallergic,  
CC antiinflammatory, and thyromimetic activity. A polynucleotide encoding a  
CC polypeptide of the invention may have a use in gene therapy. The  
CC polypeptides and polynucleotides are useful in diagnosing, treating and  
CC preventing diseases or conditions associated with the decreased  
CC expression or overexpression of ENZM, such as cell proliferative (e.g.  
CC cancer, atherosclerosis), endocrine (e.g. diabetes), neurological (e.g.  
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,  
CC allergies), developmental (e.g. Hypothyroidism, Cushing's syndrome),  
CC reproductive and vesicle-trafficking disorders, or infections. These are  
CC also useful in assessing the effects of exogenous compounds on the  
CC expression of nucleic acid and amino acid sequences of ENZM. The ENZM or  
CC its fragments are useful in screening compounds for effectiveness as  
CC agonist or antagonist of the polypeptides, or in altering the expression  
CC of the target polynucleotide and compounds that specifically bind to or  
CC modulate the activity of the polypeptide. The microarray is useful in  
CC monitoring or measuring protein-protein interactions, drug-target  
CC interactions, and gene expression profiles. The sequences shown in  
CC ADC14195-ADC14247 represent ENZM proteins of the invention.  
XX  
XX Sequence 102 AA;

Query Match 55.6%; Score 5; DB 7; Length 102;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FLRHP 9  
|||||  
Db 57 FLRHP 61

Search completed: August 31, 2006, 10:46:54  
Job time : 108.75 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocolloration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:29:54 ; Search time 139.25 Seconds  
(without alignments)  
59.786 Million cell updates/sec

Title: DENGUE\_SEROTYPE1  
Perfect score: 9  
Sequence: 1 vetwflrhp 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2849427

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : Uniprot 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	66.7	266	Q5FQA0	GLUOX
2	6	66.7	405	Q4IRY2	GIBZE
3	6	66.7	537	Q4SW56	TETNG
4	6	66.7	946	Q3P3P6	9GANM
5	6	66.7	1079	Q1133	HUMAN
6	6	66.7	1079	Q53TA5	HUMAN
7	6	66.7	1093	Q8X0R0	NEUCR
8	6	66.7	1117	Q7S795	NEUCR
9	5	55.6	29	Q6CGE3	YARLI
10	5	55.6	77	Q8A2T7	BACTN
11	5	55.6	79	Q71136	LACDL
12	5	55.6	80	Q46AP0	METBA
13	5	55.6	80	Q8TIM1	METAC
14	5	55.6	87	Q5QCQ7	CENAS
15	5	55.6	95	Q9MIS5	9TELE
16	5	55.6	101	Q3B3X5	PELID
17	5	55.6	109	Q82LF3	STRAW
18	5	55.6	115	Y115	ADE05
19	5	55.6	115	Q2K222	ADE05
20	5	55.6	118	Q3WDX0	9ACTO
21	5	55.6	118	Q4QKE8	HAIE18
22	5	55.6	122	Q5P5F8	AZOSE
23	5	55.6	124	Q2WVH8	CLOBE
24	5	55.6	124	Q47860	THEFY
25	5	55.6	126	Q4BDI1	BURVI
26	5	55.6	127	1	CRCB ERWCT
27	5	55.6	128	2	Q90ZK6 XENTH
28	5	55.6	132	2	Q6TIW1 ANETH
29	5	55.6	134	2	Q411L0 KINRA
30	5	55.6	136	2	Q977K8 9CREN
31	5	55.6	136	2	Q82HV8_STRAW

105 5 55.6 178 2 Q80A50\_9VIRU  
 106 5 55.6 178 2 Q91CC6\_9VIRU  
 107 5 55.6 178 2 Q9J0W5\_9VIRU  
 108 5 55.6 178 2 Q9J0W6\_9VIRU  
 109 5 55.6 178 2 Q9J0W7\_9VIRU  
 110 5 55.6 178 2 Q9K015\_9VIRU  
 111 5 55.6 178 2 Q10389\_9VIRU  
 112 5 55.6 178 2 Q5K003\_9VIRU  
 113 5 55.6 178 2 Q5K011\_9VIRU  
 114 5 55.6 178 2 Q5K013\_9VIRU  
 115 5 55.6 183 2 Q5WJBT\_BACSK  
 116 5 55.6 190 2 Q6L1R8\_PICTO  
 117 5 55.6 190 2 Q2UK03\_ASPOR  
 118 5 55.6 191 2 Q3QVT3\_9RHOB  
 119 5 55.6 194 2 Q98B61\_RHILO  
 120 5 55.6 196 2 Q98R28\_MYCPU  
 121 5 55.6 198 2 Q3GC97\_9FIRM  
 122 5 55.6 198 2 Q8ZP58\_SALTY  
 123 5 55.6 204 2 Q3F261\_9BURK  
 124 5 55.6 204 2 Q3G826\_9DELT  
 125 5 55.6 211 2 Q8R221\_MOUSE  
 126 5 55.6 213 2 Q4B830\_PSE14  
 127 5 55.6 213 2 Q6VEB2\_PSESY  
 128 5 55.6 215 2 Q5FJB1\_LACAC  
 129 5 55.6 218 2 Q5Z3I7\_NOCFA  
 130 5 55.6 219 2 Q47QAL\_THEFY  
 131 5 55.6 220 2 Q4HV06\_GIBZE  
 132 5 55.6 220 2 Q3W859\_9ACTO  
 133 5 55.6 220 2 Q40VL6\_KINRA  
 134 5 55.6 221 2 Q65IG8\_BACLD  
 135 5 55.6 221 2 Q9X802\_STRCO  
 136 5 55.6 224 2 Q6F256\_MESFL  
 137 5 55.6 224 2 Q9A3B2\_CAUCR  
 138 5 55.6 225 2 Q33LV9\_METHU  
 139 5 55.6 225 2 Q4H823\_9DEIO  
 140 5 55.6 231 2 Q4BWG1\_CROWT  
 141 5 55.6 232 2 Q8S5H1\_ORYSA  
 142 5 55.6 232 2 Q2T527\_BURTH  
 143 5 55.6 234 1 MENTO\_HUMAN  
 144 5 55.6 235 1 MENTO\_MOUSE  
 145 5 55.6 235 2 Q57NW5\_SALCH  
 146 5 55.6 235 2 Q5PTM7\_SALPA  
 147 5 55.6 235 2 Q8Z7C5\_SALTI  
 148 5 55.6 235 2 Q3U852\_MOUSE  
 149 5 55.6 235 2 Q3U8Q7\_MOUSE  
 150 5 55.6 235 2 Q5U205\_RAT

## ALIGNMENTS

RESULT 1  
 ID Q5FQA0\_GLUOX PRELIMINARY; PRT; 266 AA.  
 AC Q5FQA0;  
 DT 01-MAR-2005, integrated into UniProtKB/TrEMBL.  
 DT 01-MAR-2005, sequence version 1.  
 DE Putative hydrolase of the HAD superfamily.  
 GN OrderedLocusNames=GOX1706;  
 OS Gluconobacter oxydans (Gluconobacter suboxydans).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;  
 OC Acetobacteraceae; Gluconobacter.  
 OX NCBI\_TaxID=442;  
 [1]

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=621H;  
 RX PubMed=15665824; DOI=10.1038/nbr1062;  
 RA Prust C., Hoffmeister M., Liesegang H., Wierze A., Fricke W.F.,  
 RA Ehrenreich A., Gottschalk G., Deppenmeier U.,  
 RT "Complete genome sequence of the acetic acid bacterium Gluconobacter  
 oxydans.";  
 RL Nat. Biotechnol. 23:195-200(2005).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

EMBL; CP000009; AAW61446.1; -, Genomic DNA.  
 GO; GO:0016787; P:hydrolase activity; IEA.  
 GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR014200; HAD\_SF\_IIB.  
 DR InterPro; IPR006379; HAD\_SF\_IIB.  
 DR InterPro; IPR000150; Hypothet\_cof.  
 DR Pfam; PF00702; Hydrolase; 1.  
 DR TIGRFAMs; TIGR00099; Cof-subfamily; 1.  
 DR TIGRFAMs; TIGR01484; HAD-SF-IIB; 1.  
 DR PROSITE; PS01229; COF\_2; UNKNOWN\_1.  
 KW Complete proteome; Hydrolase.  
 SQ SEQUENCE 266 AA; 28559 MW; A20E08223C537EFE CRC64;

Query Match 66.7%; Score 6; DB 2; Length 266;  
 Best Local Similarity 100.0%; Pred.No. 49;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ETWFLR 7  
 Dd 98 ETWFLR 103

## RESULT 2

Q4IRY2\_GIBZE  
 ID Q4IRY2\_GIBZE PRELIMINARY; PRT; 405 AA.  
 AC Q4IRY2;  
 DT 16-AUG-2005, integrated into UniProtKB/TrEMBL.  
 DT 16-AUG-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 4.  
 DE Hypothetical protein.  
 DE ORFNames=FG00026.1;  
 OS Gibberella zeae (Fusarium graminearum).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
 OX NCBI\_TaxID=5518;  
 [1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 STRAIN=PH-1 / NRRL 31084;  
 RA Birren B.W., Nusbaum C., Abouelleil A., Allen N., Anderson S.,  
 RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,  
 RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,  
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,  
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,  
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D.,  
 RA Galagan J.E., Gardyna S., Gnerre S., Graham L., Grand-pierre N.,  
 RA Hafez N., Hagopian D., Hagos B., Hall J., Horton L., Hulme W.,  
 RA Iliev I., Jaffe D., Johnson R., Jones C., Kamal M., Kamat A.,  
 RA Karatas A., Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G.,  
 RA Lui A., Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J.,  
 RA Manning J., Matthews C., Mauceli E., McCarthy M., Meldrum J.,  
 RA Menus L., Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C.,  
 RA Nicol R., Nielsen C.B., Norbu C., O'Connor T., O'Donnell P.,  
 RA O'Neil D., Oliver J., Peterson K., Phunkhang P., Pierre N.,  
 RA Purcell S., Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C.,  
 RA Rogov P., Roman J., Schauer S., Schupbach R., Seaman S., Severy P.,  
 RA Smirnov S., Smith C., Spencer B., Stange-Thomann N., Stojanovic N.,  
 RA Stubbs M., Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,  
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,  
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,  
 RA Lander E.S.;

RT "Fusarium graminearum genome sequence.";  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

CC !- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

```

DR EMBL; AACM01000002; EAA69365.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 405 AA; 43960 MW; 5E275BB334F6AE3E CRC64;

Query Match 66.7%; Score 6; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLRH 8
DB 173 TWFLRH 178

RESULT 3
Q4SW56 TETNG PRELIMINARY; PRT; 537 AA.
AC Q4SW56;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Chromosome undetermined SCAF13690, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00011650001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lufalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bérardinis V.,
RA Cruaud C., Duprat S., Bottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet P., Schacher V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL; CAAE01013690; CAF95126.1; -; Genomic_DNA.
DR SMR; Q4SW56; 300-346.
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; Filament; 1.
DR Pfam; PF04732; Filament_head; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament.
FT NON_TER 537
FT SEQUENCE 537 AA; 60495 MW; C9178731B9AE5195 CRC64;

Query Match 66.7%; Score 6; DB 2; Length 537;
Best Local Similarity 100.0%; Pred. No. 85;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWFL 6
DB 254 VETWFL 259

RESULT 4
Q3P3P6 9GAMM PRELIMINARY; PRT; 946 AA.
AC Q3P3P6;
DT 25-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein precursor.
GN ORFNames=SdenDRAFT_1215;
OS Shewanella denitrificans OS217.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC OC Shewanellaceae; Shewanella.
OX NCBI_TaxID=318161;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OS-217;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.,
RT "Sequencing of the draft genome and assembly of Shewanella
RT denitrificans OS-217.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OS-217;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Shewanella denitrificans
RT OS-217.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL; AAU01000005; EAN71055.1; -; Genomic_DNA.
DR InterPro; IPR002035; VWF_A.
DR PRINTS; PR00453; VWFADOMAIN.
KW Hypothetical protein; Signal.
FT SIGNAL 1 31
FT SEQUENCE 946 AA; 98692 MW; EB5778488473962D CRC64;

Query Match 66.7%; Score 6; DB 2; Length 946;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWFL 6
DB 879 VETWFL 884

RESULT 5
GP113 HUMAN STANDARD; PRT; 1079 AA.
ID GP113 HUMAN STANDARD; PRT; 1079 AA.
AC Q81ZF5; Q6UXT7; Q86SL7; Q8IXD8; O8TDT3;
DT 15-FEB-2005, integrated into UniProtKB/Swiss-Prot.
DT 01-MAR-2003, sequence version 1.
DT 07-MAR-2006, entry version 25.
DE Probable G-protein coupled receptor 113 precursor (G-protein coupled
DE receptor PGR23).
GN Name=GP113; Synonyms=PGR23; ORFNames=UNQ9196/PRO34000;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

```



OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).  
RX PubMed=12435584; DOI=10.1016/S0014-5793(02)03574-3;  
RA Fredriksson R., Lagerstrom M.C., Hoeglund P.J., Schioeth H.B.;  
RT "Novel human G protein-coupled receptors with long N-terminals  
containing GPS domains and Ser/Thr-rich regions.";  
RL FEBS Lett. 531:407-414 (2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,  
RA Tautumi S., Aburatani H., Asai K., Akiyama Y.;  
RT "Genome-wide discovery and analysis of human seven transmembrane helix  
receptor genes.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2), AND VARIANT  
THR-404.  
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,  
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,  
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
RA Vandlen R.L., Watanabe C., Wiand D., Woods K., Xie M.-H., Goddard A.D.,  
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z.,  
RA Wood W.I., Godowski P.J., Gray A.M.;  
RT "The secreted protein discovery initiative (SPDI), a large-scale  
effort to identify novel human secreted and transmembrane proteins: a  
bioinformatics assessment.";  
RL Genome Res. 13:2285-2270(2003).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] OF 563-1079.  
RX MEDLINE=22040266; PubMed=12044878; DOI=10.1016/S0014-5793(02)02775-8;  
RA Takeda S., Kadowaki S., Haga T., Takaue H., Mitaku S.;  
RT "Identification of G protein-coupled receptor genes from the human  
genome sequence.";  
RL FEBS Lett. 520:97-101(2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 797-942.  
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;  
RA Vassiliatis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,  
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,  
RA Bergmann J.E., Gatanaris G.A.;  
RT "The G protein-coupled receptor repertoire of human and mouse.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).  
CC -!- FUNCTION: Orphan receptor.  
CC -!- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q8IZF5-1; Sequences=Displayed;  
CC Name=2;  
CC IsoId=Q8IZF5-2; Sequences=VSP\_012815, VSP\_012816, VSP\_012817;  
CC Note=Ref.3 (AAQ88581) sequence is in conflict in position:  
CC 888:K->N;  
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 2 family.  
CC LN-TM7 subfamily.  
CC -!- SIMILARITY: Contains 1 GPS domain.  
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous  
CC gene model prediction.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
CC EMBL: AY140955; AAN46669.1; -; mRNA.  
CC EMBL: AB056599; BAC45265.1; ALT\_SEQ; Genomic\_DNA.  
CC EMBL: AY358172; AAQ88539.1; -; mRNA.  
CC EMBL: AY358214; AAQ88581.1; -; Other\_RNA.  
CC EMBL: AB083619; BAB89332.1; ALT\_INIT; Genomic\_DNA.  
CC EMBL: AY255611; AAO85123.1; -; mRNA.

DR Ensembl; ENSG00000173567; Homo sapiens.  
DR HGNC; HGNC:18989; GPR113.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR GO; GO:0004930; F:G-protein coupled receptor activity; TAS.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; TAS.  
DR InterPro; IPR000832; GPCR\_secretin.  
DR InterPro; IPR001879; hormone\_rcpt.  
DR InterPro; IPR000203; PKD\_cys\_rich.  
DR Pfam; PF00002; 7tm\_2; 1.  
DR Pfam; PF01825; GPS; 1.  
DR PRINTS; PRO249; GPCRSECRETIN.  
DR SMART; SM00303; GPS; 1.  
DR PROSITE; PS00649; G\_PROTEIN\_RECP\_F2\_1; FALSE\_NEG.  
DR PROSITE; PS00650; G\_PROTEIN\_RECP\_F2\_2; 1.  
DR PROSITE; PS00627; G\_PROTEIN\_RECP\_F2\_3; 1.  
DR PROSITE; PS00627; G\_PROTEIN\_RECP\_F2\_3; 1.  
DR PROSITE; PS00621; G\_PROTEIN\_RECP\_F2\_4; 1.  
DR PROSITE; PS00621; GPS; 1.  
KW Alternative splicing; G-protein coupled receptor; Glycoprotein;  
KW Membrane; Polymorphism; Receptor; Signal; Transducer; Transmembrane.  
FT SIGNAL 1 25  
FT CHAIN 26 1079  
FT PROBABILE G-protein coupled receptor 113.  
FT /FTID=PRO\_0000012893.  
FT Extracellular (Potential).  
FT 1 (Potential).  
FT Cytoplasmic (Potential).  
FT 2 (Potential).  
FT Extracellular (Potential).  
FT 3 (Potential).  
FT Cytoplasmic (Potential).  
FT 4 (Potential).  
FT Extracellular (Potential).  
FT 5 (Potential).  
FT Cytoplasmic (Potential).  
FT 6 (Potential).  
FT Extracellular (Potential).  
FT 7 (Potential).  
FT Cytoplasmic (Potential).  
FT GPS.  
FT N-linked (GlcNAc. .) (Potential).  
FT N-linked (GlcNAc. .) (Potential).  
FT N-linked (GlcNAc. .) (Potential).  
FT N-linked (GlcNAc. .) (Potential).  
FT N-linked (GlcNAc. .) (Potential).  
FT N-linked (GlcNAc. .) (Potential).  
FT MVCSAPLLLLATLPLLGSPVAQASQVSETGVPRGLQ  
FT RRMQGLGRKAWNERIDRPPACPIPLSSFGWRPGQT  
FT MWAQSTLTLEEL -> MTRKLSAHSATPGYKAVTHK  
FT HHTGWARMAKTGLPEK (in isoform 2).  
FT /FTID=VSP\_012815.  
FT Missing (in isoform 2).  
FT /FTID=VSP\_012816.  
FT VSCCLQLSCAKSMSEGIWPWSSDMDGTARS -> ATNEG  
FT CILEHSGSGSTARKTDASE (in isoform 2).  
FT /FTID=VSP\_012817.  
FT A -> T (in dbSNP:2052937).  
FT /FTID=VAR\_024475.  
FT A -> V (in Ref. 3; AAQ88539).  
FT H -> Y (in Ref. 3; AAQ88539).  
FT M -> T (in Ref. 3; AAQ88581).  
FT SEQUENCE 1079 AA; 116341 MW; A18CA158F4DDBB9C CRC64;  
Query Match 66.7%; Score 6; DB 1; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 4 WFLRHP 9  
Db 267 WFLRHP 272  
RESULT 6  
Q53TAS\_HUMAN PRELIMINARY; PRT; 1079 AA.  
ID Q53TAS\_HUMAN



AC Q53TAS; 2005, integrated into UniProtKB/TrEMBL.  
 DT 24-MAY-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 3.  
 DE Hypothetical protein GPR113.  
 GN Name=GPR113;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Swearingen S., Cordes M., Cotton M.;  
 RT "The sequence of Homo sapiens BAC clone RP11-499P9.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Waterston R.H.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Waterston R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Wilson R.K.;  
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; AC010896; AAY14645.1; -; Genomic DNA.  
 DR Ensembl; ENSG00000173567; Homo sapiens.  
 DR GO; GO:0016020; C-membrane; IEA.  
 DR GO; GO:0004930; P:G-protein coupled receptor activity; IEA.  
 DR GO; GO:0007218; P:neuropeptide signaling pathway; IEA.  
 DR InterPro; IPR013032; EGF like reg.  
 DR InterPro; IPR000832; GPCR secretin.  
 DR InterPro; IPR001879; hormone rcpt.  
 DR InterPro; IPR000203; PKD\_cys\_rich.  
 DR Pfam; PF00002; 7tm\_2; 1.  
 DR Pfam; PF01825; GPS; 1.  
 DR PRINTS; PR00249; GPCRSECRETIN.  
 DR SMART; SM00303; GPS; 1.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN 1.  
 DR PROSITE; PS00650; G\_PROTEIN\_RECEP\_F2\_2; UNKNOWN 1.  
 DR PROSITE; PS0227; G\_PROTEIN\_RECEP\_F2\_3; 1.  
 DR PROSITE; PS0261; G\_PROTEIN\_RECEP\_F2\_4; 1.  
 DR PROSITE; PS50221; GPS; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1079 AA; 116341 MW; A18CA158F4DDBB9C CRC64;  
  
 Query Match 66.7%; Score 6; DB 2; Length 1079;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 4 WFLRHP 9  
 Db 267 WFLRHP 272  
  
 RESULT 7  
 QBXOR0\_NEUCR PRELIMINARY; PRT; 1093 AA.  
 AC QBXOR0;  
 DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.  
 DT 01-MAR-2002, sequence version 1.  
 DE Hypothetical protein SE6.080.  
 GN Name=SE6.080;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OC NCBI\_TaxID=5141;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA German Neurospora genome project;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; AL670004; CAD21248.1; -; Genomic DNA.  
 DR Hypothetical protein.  
 SQ SEQUENCE 1093 AA; 120695 MW; 9F6BF07A9AD661BD CRC64;  
  
 Query Match 66.7%; Score 6; DB 2; Length 1093;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 4 WFLRHP 9  
 Db 621 WFLRHP 626  
  
 RESULT 8  
 Q7S795\_NEUCR PRELIMINARY; PRT; 1117 AA.  
 AC Q7S795;  
 DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.  
 DT 15-DEC-2003, sequence version 1.  
 DT 07-FEB-2006, entry version 9.  
 DE Predicted protein.  
 GN ORFNames=NCU08869.1;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OC NCBI\_TaxID=5141;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RX MEDLINE=22598136; PubMed=12712197; DOI=10.1038/nature01554;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Ianakiev P., Bell-Pedersen D., Nelson M.A.,  
 RA Werner-Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L.,  
 RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,  
 RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,  
 RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamysseis M.,  
 RA Maucelli E., Bleke C., Rudd S., Frishman D., Krystofova S.,  
 RA Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S., Cogoni C.,  
 RA Macino G., Catchside D.E.A., Li W., Pratt R.J., Osmani S.A.,  
 RA DeSouza C.P.C., Glass N.L., Orbach M.J., Berglund J.A., Voelker R.,  
 RA Yarden O., Planann M., Seiler S., Dunlap J.C., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Eboole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.W.;  
 RT "The genome sequence of the filamentous fungus Neurospora crassa.";  
 RL Nature 422:859-868(2003).  
 CC !- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; AABX01000300; EAA31426.1; -; Genomic DNA.  
 DR SEQUENCE 1117 AA; 123343 MW; 198B80ECD607752D CRC64;  
  
 Query Match 66.7%; Score 6; DB 2; Length 1117;

```
Best Local Similarity 100.0%; Pred. No. 1.5e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 4 FLRHP 9
DB 621 WFLRHP 626

RESULT 9
Q6CGE3_YARLI PRELIMINARY; PRT; 29 AA.
AC Q6CGE3;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Similarity.
GN OrderedLocusNames=YALI0A20042g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CLIB 122 / E 150;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anchoard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykaert C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth A.,
RA Zenlou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; CR382127; CAG84207.1; -; Genomic DNA.
KW Complete proteome.
SQ SEQUENCE 29 AA; 3169 MW; 5FB2DCF7AA4626ED CRC64;

Query Match 55.6%; Score 5; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 5 FLRHP 9
DB 15 WFLRHP 19

RESULT 10
Q8A2T7_BACTN PRELIMINARY; PRT; 77 AA.
AC Q8A2T7;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Hypothetical protein.
GN OrderedLocusNames=BT3218; ORFNames=BT_3218;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AE015928; AAO78324.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 77 AA; 9164 MW; 115052AB1896BA18 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 5 FLRHP 9
DB 66 FLRHP 70

RESULT 11
Q71136_LACDL PRELIMINARY; PRT; 79 AA.
AC Q71136;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Sufi protein (Fragment).
OS Lactobacillus delbrueckii subsp. lactis.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=29397;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 4797;
RA Langenheim J.F., Ulrich R.L.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AF496414; AAQ07102.1; -; Genomic DNA.
KW Complete proteome.
SQ SEQUENCE 79 AA; 9042 MW; 5E609E1F96CFFD10 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 5 FLRHP 9
DB 44 FLRHP 48

RESULT 12
Q46AP0_METBA PRELIMINARY; PRT; 80 AA.
AC Q46AP0;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN OrderedLocusNames=Mbar_A2121;
OS Methanosarcina barkeri.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2208;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
```

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Fusaro / DSM 804;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Goodwin L.A., Saunders E.H.,
RA Schmutz J., Larimer F., Land M., Anderson I., Richardson P.,
RT "Complete sequence of chromosome 1 of Methanosaarcina barkeri str.
RT Fusaro.";
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000099; AA271052.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 80 AA; 9187 MW; 2DD0ED9A5B5C4CB3 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 5 FLRHP 9
DB 32 FLRHP 36

RESULT 13
Q8T1M1 METAC
ID Q8T1M1 METAC PRELIMINARY; PRT; 80 AA.
AC Q8T1M1
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-MAR-2006, entry version 11.
DE Hypothetical protein.
DE ORFNames=MA_4126;
OS Methanosaarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosaarcinales;
OC Methanosaarcinaceae; Methanosaarcina.
OX NCBI_TaxID=2214;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Athoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosaarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AE010299; AA007474.1; -; Genomic_DNA.
DR GenomeReviews; AE010299 GR; MA4126.
DR BioCyc; MAC188937; MA4126-MONOMER; -.
DR InterPro; IPR012933; Ycfa.
DR Pfam; PF07927; Ycfa; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 80 AA; 9184 MW; DF02721324F5D173 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 5 FLRHP 9
DB 32 FLRHP 36

```

```

DB 32 FLRHP 36

RESULT 14
Q5QCQ7 CENAS
ID Q5QCQ7 CENAS PRELIMINARY; PRT; 97 AA.
AC Q5QCQ7;
DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
DT 04-JAN-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein (fragment).
DE Hypothetical protein (fragment).
OS Cenibacterium arsenoxidans.
OC Bacteria; Cenibacterium.
OX NCBI_TaxID=204773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=UUPasi;
RA Carapito C., Muller D., Turlin E., Riegel P., Leize E., Danchin A.,
RA Van Dorsselaer A., Bertin P., Lett M.-C.;
RT "pleiotropic effect of arsenic stress on Cenibacterium arsenoxidans, a
RT metalloresistant beta-proteobacterium.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY728027; AAV68356.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON TER 87
SQ SEQUENCE 87 AA; 9653 MW; EAAD40B0A2E3C86 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 4 WFLRH 8
DB 20 WFLRH 24

RESULT 15
Q9MIS5 9TELE
ID Q9MIS5 9TELE PRELIMINARY; PRT; 95 AA.
AC Q9MIS5;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Cytochrome b (fragment).
DE Cytochrome b (fragment).
GN Names=cytb;
OS Retropinna tasmanica.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Retropinnidae; Retropinna.
OX NCBI_TaxID=89573;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22111806; PubMed=12116439; DOI=10.1080/106351500750049824;
RA Waters J.M., Lopez J.A., Wallis G.P.;
RT "Molecular phylogenetics and biogeography of galaxiid fishes
RT (Osteichthyes: Galaxiidae): dispersal, vicariance and the position of
RT Lepidogalaxias salamandroides.";
RL Syst. Biol. 49:777-795(2000).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
CC -----

```

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; AF112321; AAF67414.1; -; Genomic\_DNA.

DR SMR; Q9MIS5; 1-95.

DR DR GO; GO:0016021; C:integral to membrane; IEA.

DR DR GO; GO:0016020; C:membrane; IEA.

DR DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.

DR DR GO; GO:0005739; C:mitochondrion; IEA.

DR DR GO; GO:0005506; F:iron ion binding; IEA.

DR DR GO; GO:0046872; F:metal ion binding; IEA.

DR DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR DR GO; GO:0006118; P:electron transport; IEA.

DR DR Pfam; PF00033; Cytochrome\_B\_N\_1.

DR PROSITE; PS51002; CYTB\_NFER; 1.

KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;

KW Respiratory chain; Transmembrane; Transport.

FT NON\_TER 1 95

FT SEQUENCE 95 AA; 10578 MW; E7F5ABDD28E269DE CRC64;

QY 2 ETWFL 6

DB 77 ETWFL 81

Query Match 55.6%; Score 5; DB 2; Length 95;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETWFL 6

DB 77 ETWFL 81

RESULT 16

Q3B3X5\_PELLD PRELIMINARY; PRT; 101 AA.

AC Q3B3X5;

DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.

DT 22-NOV-2005, sequence version 1.

DT 07-FEB-2006, entry version 3.

DE Hypothetical protein.

GN ORFNames=Plut\_1094;

OS Pelodictyon luteolum (strain DSM 273) (Chlorobium luteolum (strain DSM

273)).

OC Bacteria; Chlorobii; Chlorobiales; Chlorobiaceae;

OC Chlorobium/Pelodictyon group; Pelodictyon.

OX NCBI\_TaxID=319225;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=DSM 273;

RG US DOE Joint Genome Institute;

RA Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,

RA Hammon N., Istrani S., Pitluck S., Bryant D., Schmutz J., Larimer F.,

RA Land M., Kyrpides N., Ivanova N., Richardson P.,

RT "Complete sequence of Pelodictyon luteolum DSM 273.";

RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; CP000096; AB23956.1; -; Genomic\_DNA.

KW Hypothetical protein.

SQ SEQUENCE 101 AA; 11173 MW; AAEP2D3DE11B891C CRC64;

Query Match 55.6%; Score 5; DB 2; Length 101;

Best Local Similarity 100.0%; Pred. No. 3.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7

DB 26 TWFLR 30

RESULT 17

Q82LF3\_STRAW

ID Q82LF3\_STRAW PRELIMINARY; PRT; 109 AA.  
 AC Q82LF3;  
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2003, sequence version 1.  
 DT 07-FEB-2006, entry version 14.  
 DE Hypothetical protein.  
 GN OrderedLocusNames=SAV2057;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; BA000030; BAC59768.1; -; Genomic DNA.  
 DR Biocyc; SAVE227882:SAV2057-MONOMER; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR InterPro; IPR011991; Wing\_hix\_DNA\_Bd.  
 KW Complete proteome; DNA-binding; Hypothetical protein; Transcription;  
 KW Transcription regulation.  
 SQ SEQUENCE 109 AA; 12127 MW; 9BF1F50C411DFAD2 CRC64;  
 Query Match 55.6%; Score 5; DB 2; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRHP 9  
 DB 77 FLRHP 81  
 RESULT 18  
 Y115\_ADE02  
 ID Y115\_ADE02 STANDARD; PRT; 115 AA.  
 AC P03230;  
 DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.  
 DT 21-JUL-1986, sequence version 1.  
 DT 07-FEB-2006, entry version 21.  
 DE Hypothetical protein E-115.  
 OS Human adenovirus 2 (HAdV-2).  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=10515;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
 RX MEDLINE=83056843; PubMed=7142161;  
 RA Gingeras T.R., Sciaky D., Gellinas R.E., Bing-Dong J., Yen C.E.,  
 RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;  
 RT "Nucleotide sequences from the adenovirus-2 genome.";  
 RL J. Biol. Chem. 257:13475-13491(1982).  
 CC -----

```

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; J01917; -; NOT_ANNOTATED_CDS; Genomic_DNA.
CC PIR; A03862; A03862.
CC DR Hypothetical protein.
CC FT CHAIN 1 115
CC FT SEQUENCE 115 AA; 12236 MW; C7A08EA239B8FD98 CRC64;
CC
Query Match 55.6%; Score 5; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ETWFL 6
Db 3 ETWFL 7
RESULT 19
Q2KS22_ADE05
ID Q2KS22_ADE05 PRELIMINARY; PRT; 115 AA.
AC Q2KS22;
DT 07-MAR-2006, integrated into UniProtKB/TrEMBL.
DT 07-MAR-2006, sequence version 1.
DE Hypothetical 12 kDa early protein.
OS Human adenovirus 5 (HAdV-5).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28285;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NHRC Ad5FS 7151;
RG Epidemic Outbreak Surveillance (EOS);
RA Tibbets C., Purkayastha A., Su J., Russell K., Carlisle S.,
RA Ospina R., Reynolds T., Rowley R., Hanson E., Seto D.;
RT "The complete nucleotide sequence and genome organization of Human
RT adenovirus serotype 5, field strain."
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AY601635; AAW65500.1; -; Genomic_DNA.
CC KW Hypothetical protein.
CC SEQUENCE 115 AA; 12210 MW; DF1B2DA239AA7F08 CRC64;
CC
Query Match 55.6%; Score 5; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ETWFL 6
Db 3 ETWFL 7
RESULT 20
Q3WDX0_9ACTO
ID Q3WDX0_9ACTO PRELIMINARY; PRT; 118 AA.
AC Q3WDX0;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Putative cytochrome P450.
GN OFPNAMES=FraneanlDRAFT_5182;
OS Frankia sp. EAN1pec.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Frankineae; Frankiaceae; Frankia.
OX NCBI_TaxID=298653;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EAN1pec;
RG US DOE Joint Genome Institute (JGI-PGF);

```

```

RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hamon N., Iserani S., Pitluck S., Richardson P.;
RA "Sequencing of the draft genome and assembly of Frankia sp. EAN1pec.";
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RA [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=EAN1pec;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RA "Annotation of the draft genome assembly of Frankia sp. EAN1pec.";
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AAI101000014; EAN16946.1; -; Genomic_DNA.
CC DR GO; 0020037; F:heme binding; IEA.
CC DR GO; 0005506; F:iron ion binding; IEA.
CC DR GO; 0004497; F:monooxygenase activity; IEA.
CC DR GO; 0006118; P:electron transport; IEA.
CC DR InterPro; IPR002397; BP450.
CC DR InterPro; IPR001128; Cytochrome_P450.
CC DR PRINTS; PR00359; BP450.
CC SQ SEQUENCE 118 AA; 13148 MW; 8EEA8775EPE424AD CRC64;
CC
Query Match 55.6%; Score 5; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 FLRHP 9
Db 61 FLRHP 65
RESULT 21
Q4QKE8_HAE18
ID Q4QKE8_HAE18 PRELIMINARY; PRT; 118 AA.
AC Q4QKE8;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Putative integrase/recombinase.
GN OrderedLocusNames=NTN11711;
OS Haemophilus influenzae (strain 86-028NP).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=281310;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,
RA Zhong H., Gibson J., Gibson M., Johnson L.S., Lewis L., Bakalatz L.O.,
RA Munson R.S. Jr.;
RT "Genomic sequence of an otitis media isolate of nontypeable
RT Haemophilus influenzae: comparative study with H. influenzae serotype
RT d, strain KW20."
RL J. Bacteriol. 187:4627-4636 (2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; CP000057; AAX88499.1; -; Genomic_DNA.
CC DR GO; 0003677; F:DNA binding; IEA.
CC DR GO; 0015074; P:DNA integration; IEA.
CC DR GO; 0006310; P:DNA recombination; IEA.
CC KW Complete proteome.
CC SQ SEQUENCE 118 AA; 13822 MW; 9FCB660420C82E38 CRC64;
CC
Query Match 55.6%; Score 5; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 4e+02;

```

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 21 FLRHP 25

RESULT 22
QSP5F8_AZOSE
ID QSP5F8_AZOSE PRELIMINARY; PRT; 122 AA.
AC QSP5F8;
DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein.
GN OrderedLocusNames=AZOSEA13290; ORFNames=eba2387;
OS Azoarcus sp. (strain EbN1).
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Azoarcus.
OX NCBI_TaxID=76114;
RN [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=1551059; DOI=10.1007/s00203-004-0742-9;
RA Rabus R., Kube M., Heider J., Beck A., Heitmann K., Widdel F.,
RA Reinhardt R.;
RT "The genome sequence of an anaerobic aromatic-degrading denitrifying
RT bacterium, strain EbN1.";
RL Arch. Microbiol. 183:27-36(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR555306; CA107454.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 122 AA; 13581 MW; 18B790A94ECD3255 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 16 FLRHP 20

RESULT 23
Q2WVH8_CLOBE
ID Q2WVH8_CLOBE PRELIMINARY; PRT; 124 AA.
AC Q2WVH8;
DT 10-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=CbeIDRAFT_4859;
OS Clostridium beijerinckii NCIMB 8052.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=290402;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=NCIMB 8052;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Clostridium
RT beijerinckii NCIMB 8052.";
RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=NCIMB 8052;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Clostridium beijerinckii

```

```

RT NCIMB 8052.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAL001000001; EAP62267.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 124 AA; 14913 MW; DAEBF01C68741D25 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWF 5
Db 47 VETWF 51

RESULT 24
Q47S60_THEFY
ID Q47S60_THEFY PRELIMINARY; PRT; 124 AA.
AC Q47S60;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN OrderedLocusNames=Tfu_0669;
OS Thermobifida fusca (strain YX).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptosporangineae; Nocardiopsaceae; Thermobifida.
OX NCBI_TaxID=269800;
RN [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP US DOE Joint Genome Institute;
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Di Bartolo G., Chain P., Schmutz J.,
RA Larimer F., Land M., Lykidis A., Richardson P.;
RT "Complete sequence of Thermobifida fusca YX.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000088; AA254707.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 124 AA; 14456 MW; 5F749F9A86A83FC0 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8
Db 115 WFLRH 119

RESULT 25
Q4BD11_BURVI
ID Q4BD11_BURVI PRELIMINARY; PRT; 126 AA.
AC Q4BD11;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Hypothetical protein.
GN ORFNames=Bcep1808DRAFT_1810;
OS Burkholderia vietnamiensis G4.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia;
OX NCBI_TaxID=269482;

```

```

RN NUCLEOTIDE SEQUENCE.
RC STRAIN=G4;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RG Hammon N., Israni S., Pitluck S., Richardson P.;
RA "Sequencing of the draft genome and assembly of Burkholderia
RT vietnamiensis G4.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=G4;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia vietnamiensis
RT G4.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=G4;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RG Hammon N., Israni S., Pitluck S., Richardson P.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RL "CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAEH0200037; EAM27867.1; -; Genomic_DNA.
DR TIGRFAMs; TIGR011944; CHP2246.
DR TIGRFAMs; TIGR02246; Cons_hypoth_2246; 1.
KW Hypothetical protein.
SQ SEQUENCE 126 AA; 14044 MW; A743CE0B7C1547BE CRC64;

Query Match 55.6%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWF 5
DB 12 VETWF 16

RESULT 26
RCB_ERWCT STANDARD; PRT; 127 AA.
AC Q6D7R0;
DT 05-JUL-2005, integrated into UniProtKB/Swiss-Prot.
DT 16-AUG-2004, sequence version 1.
DT 07-MAR-2006, entry version 15.
DE Protein crCB homolog.
GN Names=crCB; OrderedLocusNames=ECA1295;
OS Erwinia carotovora subsp. atroseptica (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebaihia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holve M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Trench I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).

CC -!- SUBCELLULAR LOCATION: Bacterial cell inner membrane; multi-pass
CC membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the crCB family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BX950851; CAG74205.1; -; Genomic_DNA.
DR GeneReviews; BX950851.GR; ECA1295.
DR HAMAP; MF_00454; -; 1.
DR InterPro; IPR003691; Camphor_CrCB.
DR Pfam; PF02537; CRCB; 1.
DR TIGRFAMs; TIGR00494; crCB; 1.
KW Complete proteome; Inner membrane; Membrane; Transmembrane.
FT CHAIN 1..127 /FTId=PRO_0000110100.
FT TRANSMEM 4..24 Potential.
FT TRANSMEM 35..55 Potential.
FT TRANSMEM 71..91 Potential.
FT TRANSMEM 103..123 Potential.
SQ SEQUENCE 127 AA; 13391 MW; ADED63C701397633 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 55 FLRHP 59

RESULT 27
Q90Z26_XENTR
ID Q50Z26_XENTR PRELIMINARY; PRT; 128 AA.
AC Q90Z26;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Xcat-2.
GN Names=Xcat-2;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Vempati U.D., King M.L.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF256086; AAK49295.1; -; mRNA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006445; P:regulation of translation; IEA.
DR InterPro; IPR008705; Nanos_RNA_bd.
DR Pfam; PF05741; zf-nanos; 1.
SQ SEQUENCE 128 AA; 14140 MW; E79556DEF1C0880B CRC64;

Query Match 55.6%; Score 5; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 117 FLRHP 121

RESULT 28
Q6T1W1_ANETH
ID Q6T1W1_ANETH PRELIMINARY; PRT; 132 AA.

```



```

AC O6T1W1;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 16.
DE Putative transposase.
OS Aneurinibacillus thermoaerophilus.
OC Bacteria; Firmicutes; Bacilliales; Paenibacillaceae;
OC Aneurinibacillus group; Aneurinibacillus.
OX NCBI_TaxID=143495;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=L420-91T;
RX PubMed=15044388; DOI=10.1093/glycob/cwh064;
RA Schaffer C., Messner P.;
RT "Surface-layer glycoproteins: an example for the diversity of
RT bacterial glycosylation with promising impacts on nanobiotechnology.";
RL Glycobiology 14:31R-42R(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L420-91T;
RX PubMed=15316277;
RA Novotny R., Pfoestl A., Messner P., Schaffer C.;
RT "Genetic organization of chromosomal S-layer glycan biosynthesis loci
RT of Bacillaceae.";
RL Glycoconj. J. 20:435-447(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY442352; AAS55727.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR InterPro; IPR012337; RNaseH_fold.
KW Hydrolase; Nuclease.
SQ SEQUENCE 132 AA; 15708 MW; 603062293C9D57B0 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7
DB 25 TWFLR 29

RESULT 29
Q411L0 KINRA PRELIMINARY; PRT; 134 AA.
AC Q411L0;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=KradRAFT 2276;
OS Kineococcus radiotolerans SRS30216.
OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
OC Frankineae; Kineosporiaceae; Kineococcus.
OX NCBI_TaxID=266940;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=SRS30216;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Kineococcus
RT radiotolerans SRS30216.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SRS30216;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of draft genome assembly of Kineococcus radiotolerans

```

```

RT SRS30216.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SRS30216;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAF02000024; EAM74977.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 134 AA; 15058 MW; F037BBE97A0D4676 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7
DB 41 TWFLR 45

RESULT 30
Q977K8 9CREN PRELIMINARY; PRT; 136 AA.
AC Q977K8;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Secreted protein.
OS uncultured crenarchaeote 74A4.
OC Archaea; Crenarchaeota; environmental samples;
OC marine archaeal group 1.
OX NCBI_TaxID=166279;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=1172643; PubMed=1172643; DOI=10.1128/AEM.68.1.335-345.2002;
RX Beja O., Koonin E.V., Aravind L., Taylor L.T., Seitz H., Stein J.L.,
RA Bensen D.C., Feldman R.A., Swanson R.V., Delong E.F.;
RT "Comparative Genomic Analysis of Archaeal Genotypic Variants in a
RT Single Population and in Two Different Oceanic Provinces.";
RL Appl. Environ. Microbiol. 68:335-345(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF933466; AAK96100.1; -; Genomic DNA.
SQ SEQUENCE 136 AA; 15922 MW; 852D6DD1B1626B5C CRC64;

Query Match 55.6%; Score 5; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7
DB 61 TWFLR 65

RESULT 31
Q82HV8 STRAW PRELIMINARY; PRT; 136 AA.
AC Q82HV8;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Hypothetical protein.

```



```

GN OrderedLocusNames=SAV3400;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BA000030; BAC71112.1; -; Genomic DNA.
DR BioCyc; SAV227882; SAV3400-MONOMER; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 136 AA; 14797 MW; 8A1E1A1D59C1F6F3 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 11 FLRHP 15

RESULT 32
Q3Y4H2_9BACT PRELIMINARY; PRT; 137 AA.
AC Q3Y4H2_9BACT
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Putative dissimilatory nitrite reductase (Fragment).
GN Name=nirK;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Santoro A.E., Boehm A.B., Francis C.A.;
RT "Denitrifier community composition along a nitrate and salinity
RT gradient in a coastal aquifer.";
RL Appl. Environ. Microbiol. 0:0-0(2006).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ159857; AA283971.1; -; Genomic DNA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR Pfam; PF00394; Cu-oxidase; 1.
FT NON_TER 1
FT NON_TER 137
SQ SEQUENCE 137 AA; 15234 MW; D44510A09BD948C0 CRC64;

```

```

Query Match 55.6%; Score 5; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWF 5
DB 116 VETWF 120

RESULT 33
Q6IGY3_DROME PRELIMINARY; PRT; 139 AA.
AC Q6IGY3_DROME
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE HDC04272.
GN ORFNames=HDC04272;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovyev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J.D.,
RA Faro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BK003633; DAA02331.1; -; Genomic DNA.
SQ SEQUENCE 139 AA; 14909 MW; DBA4F95D68E4045D CRC64;

Query Match 55.6%; Score 5; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 78 FLRHP 82

RESULT 34
Q4NEA0_9MICC PRELIMINARY; PRT; 139 AA.
AC Q4NEA0_9MICC
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-MAR-2006, entry version 5.
DE Similar to Glutaredoxin and related proteins.
GN ORFNames=ArthDRAFT_0289;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.

```

```
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PGF-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAHG01000023; EAL94631.1; -; Genomic DNA.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0008118; P:electron transport; IEA.
DR InterPro; IPR011915; GLRX actino.
DR InterPro; IPR012336; Thioridoxin-like fd.
DR InterPro; IPR006662; Thioridoxin.
DR InterPro; IPR006663; Thioridoxin_dom2.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRPFAM; TIGR02200; GLRX actino; 1.
DR PROSITE; PS00194; THIOREDOXIN; UNKNOWN 1.
SQ SEQUENCE 139 AA; 14842 MW; 189B6C4668B8B6E6 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 19 FLRHP 23

RESULT 35
Q620F6 ORYSA PRELIMINARY; PRT; 140 AA.
AC Q620F6;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein OSJNBa0062G05.18.
GN Names:OSJNBa0062G05.18;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade;
OC Eriactoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
clone:OSJNBa0062G05.18";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF003491; BAD03650.1; -; Genomic DNA.
DR Gramene; Q620F6; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:000355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003822; PAH.
DR Pfam; PF02671; PAH; 1.
DR Hypothetical protein.
KW SEQUENCE 140 AA; 16304 MW; 2476A657C1C2FC74 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 84 FLRHP 88
```

```
RESULT 36
Q3Y4M6_9BACT PRELIMINARY; PRT; 140 AA.
AC Q3Y4M6;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Putative dissimilatory nitrite reductase (Fragment).
GN Name:nirK;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Santoro A.E., Boehm A.B., Francis C.A.;
RT "Denitrifier community composition along a nitrate and salinity
gradient in a coastal aquifer.";
RL Appl. Environ. Microbiol. 0:0-0(2006).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ159803; AAZ83917.1; -; Genomic DNA.
DR EMBL; DQ159738; AAZ83854.1; -; Genomic DNA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR Pfam; PF00394; Cu-oxidase; 1.
FT NON_TER 1
FT NON_TER 140
SQ SEQUENCE 140 AA; 15544 MW; 3D1C65CE173B6925 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWF 5
DB 119 VETWF 123

RESULT 37
Q3Y4P7_9BACT PRELIMINARY; PRT; 140 AA.
AC Q3Y4P7;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Putative dissimilatory nitrite reductase (Fragment).
GN Name:nirK;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Santoro A.E., Boehm A.B., Francis C.A.;
RT "Denitrifier community composition along a nitrate and salinity
gradient in a coastal aquifer.";
RL Appl. Environ. Microbiol. 0:0-0(2006).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ159782; AAZ83896.1; -; Genomic DNA.
DR EMBL; DQ159790; AAZ83904.1; -; Genomic DNA.
DR EMBL; DQ159750; AAZ83866.1; -; Genomic DNA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR Pfam; PF00394; Cu-oxidase; 1.
FT NON_TER 1
FT NON_TER 140
SQ SEQUENCE 140 AA; 15562 MW; 69E86F8DD32AD297 CRC64;
```

```

Query Match          55.6%; Score 5; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VETWF 5
Db      119 VETWF 123

RESULT 38
Q3Y4P8_9BACT
ID Q3Y4P8_9BACT PRELIMINARY; PRT; 140 AA.
AC Q3Y4P8;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Putative dissimilatory nitrite reductase (Fragment).
GN Name=nirk;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Santoro A.E., Boehm A.B., Francis C.A.;
RT "Denitrifier community composition along a nitrate and salinity
   gradient in a coastal aquifer.";
RL Appl. Environ. Microbiol. 0:0-0(2006).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ159781; AAZ83895.1; -; Genomic DNA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR Pfam; PF00394; Cu-oxidase; 1.
FT NON_TER 1
FT NON_TER 140
FT NON_TER 140
SQ SEQUENCE 140 AA; 15584 MW; 3865D54F75B1683A CRC64;

Query Match          55.6%; Score 5; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VETWF 5
Db      119 VETWF 123

RESULT 39
Q3Y4Q3_9BACT
ID Q3Y4Q3_9BACT PRELIMINARY; PRT; 140 AA.
AC Q3Y4Q3;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Putative dissimilatory nitrite reductase (Fragment).
GN Name=nirk;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Santoro A.E., Boehm A.B., Francis C.A.;
RT "Denitrifier community composition along a nitrate and salinity
   gradient in a coastal aquifer.";
RL Appl. Environ. Microbiol. 0:0-0(2006).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ159776; AAZ83890.1; -; Genomic DNA.
DR GO; GO:0005507; F:copper ion binding; IEA.

```

```

DR InterPro; IPR001117; Cu-oxidase.
DR Pfam; PF00394; Cu-oxidase; 1.
FT NON_TER 1
FT NON_TER 140
FT NON_TER 140
SQ SEQUENCE 140 AA; 15616 MW; 3CC71550667E6A6A CRC64;

Query Match          55.6%; Score 5; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VETWF 5
Db      119 VETWF 123

RESULT 40
Q3Y4Q4_9BACT
ID Q3Y4Q4_9BACT PRELIMINARY; PRT; 140 AA.
AC Q3Y4Q4;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Putative dissimilatory nitrite reductase (Fragment).
GN Name=nirk;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Santoro A.E., Boehm A.B., Francis C.A.;
RT "Denitrifier community composition along a nitrate and salinity
   gradient in a coastal aquifer.";
RL Appl. Environ. Microbiol. 0:0-0(2006).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ159775; AAZ83889.1; -; Genomic DNA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR Pfam; PF00394; Cu-oxidase; 1.
FT NON_TER 1
FT NON_TER 140
FT NON_TER 140
SQ SEQUENCE 140 AA; 15547 MW; 88AD7BD3C82B7628 CRC64;

Query Match          55.6%; Score 5; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VETWF 5
Db      119 VETWF 123

Search completed: August 31, 2006, 10:39:40
Job time : 142.25 secs

```

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:40:05 ; Search time 17.25 Seconds  
(without alignments)  
50.200 Million cell updates/sec

Title: DENGUE\_SEROTYPE1

Perfect score: 9

Sequence: 1 vetwflrhp 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 1

Total number of hits satisfying chosen parameters: 283347

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database :

PIR 80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	55.6	115	2 A03862	hypothetical prote
2	5	55.6	132	2 S14077	Ig kappa chain - A
3	5	55.6	168	2 B88102	protein W09G10.5 [
4	5	55.6	170	2 H69850	mutator Mutr prote
5	5	55.6	178	2 JQ1547	stripe disease-spe
6	5	55.6	196	2 F90534	transcription anti
7	5	55.6	208	2 T33341	hypothetical prote
8	5	55.6	221	2 T35525	probable two compo
9	5	55.6	224	2 B87657	conserved hypothet
10	5	55.6	225	2 C88939	protein C05E4.8 [l
11	5	55.6	235	2 AF0656	conserved hypothet
12	5	55.6	241	2 T27636	hypothetical prote
13	5	55.6	244	2 AI2644	flagellar basal bo
14	5	55.6	244	2 A97427	flgF protein (U951
15	5	55.6	247	2 S75903	hypothetical prote
16	5	55.6	270	2 E64924	hypothetical prote
17	5	55.6	270	2 D85774	hypothetical prote
18	5	55.6	270	2 H90925	hypothetical prote
19	5	55.6	296	2 E91027	hypothetical prote
20	5	55.6	296	2 F85871	hypothetical prote
21	5	55.6	296	2 G65002	hypothetical prote
22	5	55.6	300	2 T32681	hypothetical prote
23	5	55.6	304	2 D64122	hypothetical prote
24	5	55.6	313	2 AH0966	conserved hypothet
25	5	55.6	317	2 S48036	hypothetical prote
26	5	55.6	319	2 T27118	hypothetical prote
27	5	55.6	321	2 AE1068	probable membrane
28	5	55.6	343	2 H64491	hypothetical prote
29	5	55.6	347	2 A46567	tetracycline resis

30	5	55.6	358	2 T34382	hypothetical prote
31	5	55.6	368	2 T06460	anthranilate phosph
32	5	55.6	375	1 E64593	2-oxoacid-ferredox
33	5	55.6	375	2 G71919	chain of 2-oxoglut
34	5	55.6	378	2 H96773	hypothetical prote
35	5	55.6	396	2 I58168	growth factor arg3
36	5	55.6	415	2 AC3235	nitrotriacetate
37	5	55.6	415	2 B86434	protein T17H7.13 [
38	5	55.6	431	2 T01557	hypothetical prote
39	5	55.6	445	2 I38027	MLN 64 protein - h
40	5	55.6	456	2 T06589	3-methyl-2-oxobuta
41	5	55.6	473	2 JC4313	keratin 16, type I
42	5	55.6	540	2 T20352	hypothetical prote
43	5	55.6	601	2 F64116	endopeptidase Ia h
44	5	55.6	720	2 S75935	hypothetical prote
45	5	55.6	759	2 B83474	probable type II s
46	5	55.6	783	2 E86254	hypothetical prote
47	5	55.6	818	2 F82173	collagenase VCI650
48	5	55.6	905	2 I49499	alpha N-catenin I
49	5	55.6	906	2 A43000	alpha N-catenin -
50	5	55.6	945	1 A45011	hypothetical prote
51	5	55.6	946	2 T31488	alpha N-catenin II
52	5	55.6	953	2 I49500	alpha N-catenin II
53	5	55.6	1011	1 A45598	H+-exporting ATPas
54	5	55.6	1155	2 B96761	probable protein k
55	5	55.6	1841	2 T38091	cell division cont
56	4	44.4	20	2 S21176	testosterone beta
57	4	44.4	41	2 D82691	hypothetical prote
58	4	44.4	41	2 D96009	probable plasmid s
59	4	44.4	45	2 JH0208	hypothetical 5.2K
60	4	44.4	50	2 A69055	hypothetical prote
61	4	44.4	53	2 D82612	hypothetical prote
62	4	44.4	55	2 A69152	DNA-dependent RNA
63	4	44.4	61	2 F96005	hypothetical prote
64	4	44.4	63	2 T15583	hypothetical prote
65	4	44.4	65	2 D87622	hypothetical prote
66	4	44.4	67	2 S08458	hypothetical prote
67	4	44.4	69	2 A71084	hypothetical prote
68	4	44.4	72	2 G71355	probable ribosomal
69	4	44.4	73	2 A90885	hypothetical prote
70	4	44.4	73	2 F85733	hypothetical prote
71	4	44.4	74	2 T44088	probable transposa
72	4	44.4	78	2 D29653	hypothetical prote
73	4	44.4	80	2 AF2836	hypothetical prote
74	4	44.4	81	2 B84095	hypothetical prote
75	4	44.4	82	2 H64896	probable membrane
76	4	44.4	85	2 G63176	hypothetical prote
77	4	44.4	88	2 AF0549	conserved hypothet
78	4	44.4	89	2 T42967	hypothetical prote
79	4	44.4	89	2 T50245	hypothetical prote
80	4	44.4	89	2 A59100	hypothetical prote
81	4	44.4	94	2 D64446	hypothetical prote
82	4	44.4	94	2 T29563	hypothetical prote
83	4	44.4	95	2 T18160	hypothetical prote
84	4	44.4	97	2 AI0538	hypothetical prote
85	4	44.4	97	2 D95328	hypothetical prote
86	4	44.4	98	2 I49562	alpha-1 type III c
87	4	44.4	98	2 T17924	hypothetical prote
88	4	44.4	99	2 B90063	hypothetical prote
89	4	44.4	102	2 S37929	hypothetical prote
90	4	44.4	102	2 AI2711	hypothetical prote
91	4	44.4	102	2 G74993	hypothetical prote
92	4	44.4	103	2 S64330	probable membrane
93	4	44.4	105	2 A72735	hypothetical prote
94	4	44.4	105	2 G72572	hypothetical prote
95	4	44.4	106	2 T12684	hypothetical prote
96	4	44.4	106	2 S51046	hypothetical prote
97	4	44.4	106	2 E90062	hypothetical prote
98	4	44.4	109	2 T29627	hypothetical prote
99	4	44.4	110	2 S64948	probable membrane
100	4	44.4	112	2 T02744	ubiquitin conjugat
101	4	44.4	112	2 AD3596	hypothetical cytos
102	4	44.4	114	1 JQ2242	thioredoxin h - Ar

103 4 44.4 114 2 D72665 hypothetical prote  
104 4 44.4 115 2 C70074 hypothetical prote  
105 4 44.4 115 2 C72568 hypothetical prote  
106 4 44.4 115 2 F72779 hypothetical prote  
107 4 44.4 116 2 S70038 hypothetical prote  
108 4 44.4 116 2 H95414 hypothetical prote  
109 4 44.4 118 2 G95121 Tn5252, Oxf 10 pro  
110 4 44.4 120 2 C86882 hypothetical prote  
111 4 44.4 121 2 S74554 hypothetical prote  
112 4 44.4 121 2 S76514 hypothetical prote  
113 4 44.4 122 2 S32630 ribonucleoside-dip  
114 4 44.4 122 2 D72756 hypothetical prote  
115 4 44.4 123 2 H90236 conserved hypotet  
116 4 44.4 123 2 C84412 hypothetical prote  
117 4 44.4 123 2 C59822 hypothetical prote  
118 4 44.4 123 2 A27077 conserved hypotet  
119 4 44.4 124 2 S03521 Ig kappa chain pre  
120 4 44.4 124 2 C75359 hypothetical prote  
121 4 44.4 125 2 F82834 hypothetical prote  
122 4 44.4 125 2 S76468 hypothetical prote  
123 4 44.4 128 2 AF2143 hypothetical prote  
124 4 44.4 129 2 D75346 hypothetical prote  
125 4 44.4 130 2 A41911 oxoglutarate dehyd  
126 4 44.4 132 2 T50389 homolog to yeast PK  
127 4 44.4 133 2 B44370 probable G-protein  
128 4 44.4 134 2 F85589 hypothetical prote  
129 4 44.4 134 2 D90739 hypothetical prote  
130 4 44.4 134 2 F64817 probable membrane  
131 4 44.4 134 2 C84385 hypothetical prote  
132 4 44.4 135 2 S31682 inhibin beta-A cha  
133 4 44.4 135 2 C69265 hypothetical prote  
134 4 44.4 136 2 T18052 DEAH box protein a  
135 4 44.4 136 2 H72633 hypothetical prote  
136 4 44.4 136 2 JQ0421 biastocidin S-acet  
137 4 44.4 137 2 F29380 Ig heavy chain pre  
138 4 44.4 137 2 A71308 hypothetical prote  
139 4 44.4 137 2 T22872 hypothetical prote  
140 4 44.4 137 2 S40760 hypothetical prote  
141 4 44.4 141 1 H4L2C hemoglobin alpha-1  
142 4 44.4 141 2 JT0624 hemoglobin alpha 1  
143 4 44.4 141 2 T46427 hypothetical prote  
144 4 44.4 142 2 F82239 probable transcrip  
145 4 44.4 143 2 A10838 probable membrane  
146 4 44.4 146 2 T35484 hypothetical prote  
147 4 44.4 146 2 B71430 hypothetical prote  
148 4 44.4 147 2 AF0799 probable sugar pro  
149 4 44.4 147 2 C90094 hypothetical prote  
150 4 44.4 149 2 S48927 hypothetical prote

## ALIGNMENTS

RESULT 1  
A03862  
hypothetical protein E-115 - human adenovirus 2  
C:Species: Mastadenovirus h2 (human adenovirus 2)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C:Accession: A03862  
R:Gingras, T.R.; Sciaky, D.; Gelinias, R.E.; Bing-Dong, J.; Yen, C.E.; Kelly, M.M.; Bull  
J. Biol. Chem. 257, 13475-13491, 1982  
A:Title: Nucleotide sequences from the adenovirus-2 genome.  
A:Reference number: A92351; MUID:83056843; PMID:7142161  
A:Accession: A03862  
A:Molecule type: DNA  
A:Residues: 1-115 <GIN>  
A:Cross-references: UNIPROT:P03290; UNIPARC:UPI00001392B5

Query Match 55.6%; Score 5; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETWFL 6

Db 3 ETWFL 7

## RESULT 2

S14077  
Ig kappa chain - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999  
C:Accession: S14077  
R:Schwager, J.; Buerckert, N.; Schwager, M.; Wilson, M.  
EMBO J. 10, 505-511, 1991  
A:Title: Evolution of immunoglobulin light chain genes: analysis of Xenopus Igl isotypes  
A:Reference number: S14076; MUID:91160503; PMID:1705882  
A:Accession: S14077  
A:Molecule type: mRNA  
A:Residues: 1-132 <SCH>  
A:Cross-references: UNIPARC:UPI000017698D  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 55.6%; Score 5; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
Db 48 TWFLR 52

## RESULT 3

B88102  
protein W09G10.5 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: B88102  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A>Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_eleg  
A:Accession: B88102  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-168 <STO>  
A:Cross-references: UNIPROT:O16641; UNIPARC:UPI0000075172; GB:chr\_II; PIDN:AB66113.1; PII  
C:Genetics:  
A:Gene: W09G10.5  
A:Map position: 2  
C:Superfamily: Caenorhabditis elegans hypothetical protein C31G12.2

Query Match 55.6%; Score 5; DB 2; Length 168;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
Db 30 TWFLR 34

## RESULT 4

H69850  
mutator MutT protein homolog yjhB - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: H69850  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallerc  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;

C;Accession: B87657

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A>Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: AB7249; MUID:21173698; PMID:11259647  
A:Accession: B87657  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-224 <STO>  
A:Cross-references: UNIPROT:O9A3B2; UNIPARC:UPI00000C7A0F; GB:AE005673; NID:gl13424986; F  
C:Genetics:  
A:Gene: CC3292

Query Match 55.6%; Score 5; DB 2; Length 224;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
|||||  
Db 97 WFLRH 101

RESULT 10  
C88939  
protein C05E4.8 [imported] - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: C88939  
R;anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A>Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: C88939  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-225 <STO>  
A:Cross-references: UNIPROT:O17356; UNIPARC:UPI0000082F2F; GB:chr\_V; PIDN:AAE71277.1; PI  
C:Genetics:  
A:Gene: C05E4.8  
A:Map position: 5  
C:Superfamily: *Caenorhabditis* transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 184 FLRHP 188

RESULT 11  
AF0656  
conserved hypothetical protein STY1354 [imported] - *Salmonella enterica* subsp. enterica  
C:Species: *Salmonella enterica* subsp. enterica serovar typhi  
A>Note: this species has also been called *Salmonella typhi*  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AF0656  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulle, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A>Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AF0656  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-235 <PAR>

A:Cross-references: UNIPARC:UPI0000059F2E; GB:AL513382; PIDN:CAD01623.1; PID:gl6502477; C  
C:Genetics:  
A:Gene: STY1354

Query Match 55.6%; Score 5; DB 2; Length 235;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
|||||  
Db 65 WFLRH 69

RESULT 12  
T27636  
hypothetical protein ZC64.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T27636  
R;Bentley, D.  
submitted to the EMBL Data Library, October 1995  
A:Description: The sequence of *C. elegans* cosmid ZC64.  
A:Reference number: Z20397  
A:Accession: T27636  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-241 <BEN>  
A:Cross-references: UNIPROT:Q23379; UNIPARC:UPI0000081BA9; EMBL:U39740; PIDN:AAA80427.1;  
C:Genetics:  
A:Gene: CESP:ZC64.1  
C:Superfamily: *Caenorhabditis* transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 241;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 204 FLRHP 208

RESULT 13  
AI2644  
flagellar basal body rod protein [imported] - *Agrobacterium tumefaciens* (strain C58, Dup  
C:Species: *Agrobacterium tumefaciens*  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 05-Oct-2004  
C:Accession: AI2644  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E  
ster, E.W.  
A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AI2644  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-244 <KUR>  
A:Cross-references: UNIPROT:O34170; UNIPARC:UPI00000D1464; GB:AE008688; PIDN:AAL41575.1;  
A:Experimental source: strain C58 (dupont)  
C:Genetics:  
A:Gene: flgF  
A:Map position: circular chromosome  
C:Superfamily: rod protein flgF

Query Match 55.6%; Score 5; DB 2; Length 244;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 168 FLRHP 172



RESULT 14  
A97427  
flgF protein (U95165) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
C:Species: Agrobacterium tumefaciens  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 05-Oct-2004  
C:Accession: A97427  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: A97427  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-244 <KUR>  
A:Cross-references: UNIPROT:O34170; UNIPARC:UPI00000D1464; GB:AE007869; PIDN:AAK86370.1;  
C:Genetics:  
A:Gene: AGR\_C\_982  
A:Map position: circular chromosome  
C:Superfamily: rod protein flgF

Query Match 55.6%; Score 5; DB 2; Length 244;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 168 FLRHP 172  
|||||

RESULT 15  
S75903  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S75903  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, N. K.; Science 277, 129-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S75903  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-247 <KAN>  
A:Cross-references: UNIPROT:P74268; UNIPARC:UPI00001290CE; EMBL:D90913; GB:AB001339; NID  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 55.6%; Score 5; DB 2; Length 247;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 209 FLRHP 213  
|||||

RESULT 16  
E64924  
hypothetical protein b1669 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C:Accession: E64924  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.; Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: E64924

A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-270 <BLAT>  
A:Cross-references: UNIPROT:P77147; UNIPARC:UPI000013A9BC; GB:AE000262; GB:U00096; NID:91  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRH 8  
Db 20 WFLRH 24  
|||||

RESULT 17  
D85774  
hypothetical protein Z2696 [imported] - Escherichia coli (strain O157:H7, substrain EDL93;  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: D85774  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: D85774  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-270 <STO>  
A:Cross-references: UNIPROT:Q8X618; UNIPARC:UPI00000D0C14; GB:AE005174; NID:GI2515668; P  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z2696  
C:Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRH 8  
Db 20 WFLRH 24  
|||||

RESULT 18  
H90925  
hypothetical protein ECs2376 [imported] - Escherichia coli (strain O157:H7, substrain RIN  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: H90925  
R:Hayaashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: H90925  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-270 <HAY>  
A:Cross-references: UNIPROT:Q8X618; UNIPARC:UPI00000D0C14; GB:BA000007; PIDN:BA035799.1;  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECs2376  
C:Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRH 8  
Db 20 WFLRH 24  
|||||

```
Db          20 WFLRH 24

RESULT 19
hypothetical protein ECg3189 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 31-Dec-2004
C:Accession: E91027
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ohnishi, M.; Kurokawa, K.; Han, C.G.
Gisawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E91027
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <STO>
A:Cross-references: UNIPROT:Q8XCT0; UNIPARC:UPI00000D0433; GB:BA000007; PIDN:BA036612.1;
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
C:Superfamily: human PML-1 protein

Query Match          55.6%; Score 5; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
Db          17 FLRHP 21

RESULT 20
hypothetical protein yfci [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
C:Accession: F85871
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85871
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <STO>
A:Cross-references: UNIPROT:Q8XCT0; UNIPARC:UPI00000D0433; GB:AE005174; NID:g12516661; E
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yfci
C:Superfamily: human PML-1 protein

Query Match          55.6%; Score 5; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
Db          17 FLRHP 21

RESULT 21
hypothetical protein b2305 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 31-Dec-2004
C:Accession: G65002
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
```

```
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G65002
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-296 <BLAT>
A:Cross-references: UNIPROT:P77768; UNIPARC:UPI0000047C8B; GB:AE000319; GB:U00096; NID:
A:Experimental source: strain K-12, substrain MGI655
C:Superfamily: human PML-1 protein

Query Match          55.6%; Score 5; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
Db          17 FLRHP 21

RESULT 22
hypothetical protein K07C6.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32681
R:Wagner-McPherson, C.; Gilliam, B.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid K07C6.
A:Reference number: Z21209
A:Accession: T32681
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-300 <WAG>
A:Cross-references: UNIPROT:O44645; UNIPARC:UPI0000079677; EMBL:AF039049; PIDN:AAB94256.1
A:Experimental source: strain Bristol N2; clone K07C6
C:Genetics:
A:Gene: CESP:K07C6.14
A:Map position: 5
C:Superfamily: Caenorhabditis transposon Tc1 hypothetical 32K protein

Query Match          55.6%; Score 5; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
Db          231 FLRHP 235

RESULT 23
hypothetical protein H1424 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: D64122
R:Flaischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: D64122
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-304 <TIGR>
A:Cross-references: UNIPROT:P45198; UNIPARC:UPI000013AAB2; GB:U32821; GB:L42023; NID:g15

Query Match          55.6%; Score 5; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
```

Db 60 FLRHP 64

RESULT 24  
AH0966

conserved hypothetical protein STY4020 [imported] - Salmonella enterica subsp. enterica  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 31-Dec-2004  
C;Accession: AH0966  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AH0966  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-313 <PAR>  
A;Cross-references: UNIPARC:UPI000005A6D2; GB:AL513382; PIDN:CAD03228.1; PID:g16504856;  
C;Genetics:  
A;Gene: STY4020  
C;Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 313;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 17 FLRHP 21

RESULT 25  
S48036

hypothetical protein - kiwi fruit  
C;Species: Actinidia chinensis var. deliciosa (kiwi fruit)  
C;Date: 26-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 17-Mar-1999  
C;Accession: S48036  
R;Ledger, S.E.; Gardner, R.C.  
Plant Mol. Biol. 25, 877-886, 1994  
A;Title: Cloning and characterization of five cDNAs for genes differentially expressed d  
A;Reference number: S48035; MUID:94355660; PMID:8075403  
A;Accession: S48036  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-317 <LED>  
A;Cross-references: UNIPARC:UPI000012DB29; EMBL:L27809; NID:g450236; PID:g450237  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994

Query Match 55.6%; Score 5; DB 2; Length 317;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 39 FLRHP 43

RESULT 26  
T27118

hypothetical protein Y53C10A.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T27118  
R;White, S.  
Submitted to the EMBL Data Library, November 1998  
A;Reference number: Z20314  
A;Accession: T27118  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

A;Residues: 1-319 <WIL>  
A;Cross-references: UNIPROT:Q9XW52; UNIPARC:UPI0000076E2F; EMBL:AL033536; PIDN:CAA22139.1  
A;Experimental source: clone Y53C10A  
C;Genetics:  
A;Gene: CESP:Y53C10A.5  
A;Introns: 31/2; 91/1; 124/1; 233/3; 264/3; 290/1

Query Match 55.6%; Score 5; DB 2; Length 319;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VETWF 5  
Db 154 VETWF 158

RESULT 27  
AE1068

probable membrane protein STY4875 [imported] - Salmonella enterica subsp. enterica serovar  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 31-Dec-2004  
C;Accession: AE1068  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AE1068  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-321 <PAR>  
A;Cross-references: UNIPARC:UPI000005A9BD; GB:AL513382; PIDN:CAD03364.1; PID:g16505636;  
C;Genetics:  
A;Gene: STY4875  
C;Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 321;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 17 FLRHP 21

RESULT 28  
H64491

hypothetical protein MJ1537 - Methanococcus jannaschii  
C;Species: Methanococcus jannaschii  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C;Accession: H64491  
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, C.;  
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.  
A;Reference number: A64300; MUID:96337999; PMID:8688087  
A;Accession: H64491  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-343 <BUL>  
A;Cross-references: UNIPROT:Q58932; UNIPARC:UPI000013AD2A; GB:U67594; GB:L77117; NID:g15;  
C;Genetics:  
A;Map position: REV1515744-1514713  
C;Superfamily: Methanococcus jannaschii hypothetical protein MJ1537

Query Match 55.6%; Score 5; DB 2; Length 343;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 Db 105 FLRHP 109

RESULT 29  
 A46567  
 tetracycline resistance protein - Streptomyces rimosus  
 C:Species: Streptomyces rimosus  
 C:Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 09-Jul-2004  
 C:Accession: A46567  
 R:Reynes, J.P.; Calmels, T.; Drocourt, D.; Tiraby, G.  
 J. Gen. Microbiol. 134, 585-598, 1988  
 A:Title: Cloning, expression in Escherichia coli and nucleotide sequence of a tetracycline resistance gene from Streptomyces rimosus  
 A:Reference number: A46567; MUID:89036114; PMID:3053973  
 A:Accession: A46567  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-347 <REV>  
 A:Cross-references: UNIPROT:P14551; UNIPARC:UPI0000136B1D; GB:M20370; NID:g153503; PIDN:

Query Match 55.6%; Score 5; DB 2; Length 347;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
 Db 313 TWFLR 317

RESULT 30  
 T34382  
 hypothetical protein T25G12.9 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T34382  
 R:Du, Z.  
 submitted to the EMBL Data Library, December 1995  
 A:Description: The sequence of C. elegans cosmid T25G12.  
 A:Reference number: Z21515  
 A:Accession: T34382  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-358 <DUZ>  
 A:Cross-references: UNIPROT:Q22789; UNIPARC:UPI0000081C10; EMBL:U43283; PIDN:AAC69023.1;  
 A:Experimental source: strain Bristol N2; clone T25G12  
 C:Genetics:  
 A:Gene: CESP:T25G12.9  
 A:Map position: X  
 C:Superfamily: Caenorhabditis transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 358;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 Db 231 FLRHP 235

RESULT 31  
 T06460  
 anthranilate phosphoribosyltransferase (EC 2.4.2.18) - garden pea (fragment)  
 N:Alternate names: phosphoribosylanthranilate transferase  
 C:Species: Pisum sativum (garden pea)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C:Accession: T06460  
 R:Sato, N.; Kazuno, A.; Ohta, N.; Onshima, K.  
 submitted to the EMBL Data Library, June 1996  
 A:Description: Isolation of a pea cDNA for phosphoribosylanthranilate transferase.  
 A:Reference number: Z15694  
 A:Accession: T06460

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-368 <SAT>

A:Cross-references: UNIPROT:Q43085; UNIPARC:UPI00000A9D1C; EMBL:D86180; PIDN:BAAL13032.1

A:Experimental source: var. Alaska

C:Genetics:

-A:Gene: PAT1

C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 55.6%; Score 5; DB 2; Length 368;

Best Local Similarity 100.0%; Pred. No. 71;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9

Db 339 FLRHP 343

RESULT 32

E64593

2-oxoacid-ferredoxin oxidoreductase (EC 1.2.7.-) alpha chain - Helicobacter pylori (strain N1)

N:Alternate names: 2-oxoacid:ferredoxin oxidoreductase (CoA-acetylating)

C:Species: Helicobacter pylori

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: E64593

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne-

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N.

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: E64593

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-375 <TON>

A:Cross-references: UNIPROT:Q25311; UNIPARC:UPI00000D30AB; GB:AE000572; GB:AE000511; NID:

C:Superfamily: Helicobacter pylori 2-oxoacid ferredoxin oxidoreductase; 2-oxoacid ferred-

C:Keywords: oxidoreductase

F:5-186/Domain: 2-oxoacid ferredoxin oxidoreductase homology <FEO>

Query Match 55.6%; Score 5; DB 1; Length 375;

Best Local Similarity 100.0%; Pred. No. 72;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9

Db 124 FLRHP 128

RESULT 33

G71919

chain of 2-oxoglutarate oxidoreductase - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

C:Accession: G71919

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho-

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: G71919

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <ARN>

A:Cross-references: UNIPROT:Q9ZLP1; UNIPARC:UPI00000D364E; GB:AE001486; GB:AE001439; NID:

A:Experimental source: strain J99

C:Genetics:

A:Gene: oora

C:Superfamily: Helicobacter pylori 2-oxoacid ferredoxin oxidoreductase; 2-oxoacid ferred-

F:5-186/Domain: 2-oxoacid ferredoxin oxidoreductase homology <FEO>



QY 4 WFLRH 8  
|||||  
Db 65 WFLRH 69

## RESULT 38

T01557

Hypothetical protein A\_TM018A10.5 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C&gt;Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 09-Jul-2004

C:Accession: T01557

R:Dempsey, S.; Harper, M.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of A. thaliana TM018A10.

A:Reference number: Z14348

A:Accession: T01557

A:Status: translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 1-431 &lt;DEM&gt;

A:Cross-references: UNIPROT:O23088; UNIPARC:UPI000009FF8; EMBL:AF013294; NID:g2252848;

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Note: A\_TM018A10.5

Query Match 55.6%; Score 5; DB 2; Length 431;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWF 5  
|||||  
Db 152 VETWF 156

## RESULT 39

I38027

MLN 64 protein - human

C:Species: Homo sapiens (man)

C&gt;Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 09-Jul-2004

C:Accession: I38027; S60682

R:Tomasetto, C.; Regnier, C.H.; Moog-Lutz, C.; Mattei, M.G.; Chenard, M.P.; Lidereau, R.

Genomics 28, 367-376, 1995

A:Title: Identification of four novel human genes amplified and overexpressed in breast

A:Reference number: I37080; MUID:96039245; PMID:7490069

A:Accession: I38027

A:Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: mRNA

A:Residues: 1-445 &lt;RES&gt;

A:Cross-references: UNIPROT:Q14849; UNIPARC:UPI000012F1BC; EMBL:X80198; NID:g951278; PID

A:Note: submitted to the EMBL Data Library, July 1994

C:Genetics:

A:Gene: MLN64

Query Match 55.6%; Score 5; DB 2; Length 445;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETWFL 6  
|||||  
Db 165 ETWFL 169

## RESULT 40

T06589

3-methyl-2-oxobutanoate dehydrogenase (lipoamide) (EC 1.2.4.4) El-alpha chain precursor,

N;Alternate names: branched-chain alpha-keto acid dehydrogenase

C:Species: Lycopersicon esculentum (tomato)

C&gt;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T06589

R:Glritch, A.; Baumelein, H.

submitted to the EMBL Data Library, April 1997

A:Description: A molecular cloning and characterization of cDNA coding for the branched

A:Reference number: Z15779

A:Accession: T06589  
A:Status: preliminary; translated from GB/EMBL/DBBJ  
A:Molecule type: mRNA  
A:Residues: 1-456 <GIR>

A:Cross-references: UNIPROT:O03849; UNIPARC:UPI00000A796F; EMBL:Z94180; PIDN:CAB08111.1

A:Experimental source: cultivar Bonner Beste, mutant chloronerva; roots

C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin

C:Keywords: mitochondrion; oxidoreductase; phosphoprotein

F:1-14/Domain: transit peptide (mitochondrion) #status predicted &lt;TNP&gt;

F:35-456/Product: 3-methyl-2-oxobutanoate dehydrogenase (lipoamide) El-alpha chain #statu

F:239-286/Domain: thiamin pyrophosphate-binding domain homology &lt;TPB&gt;

Query Match 55.6%; Score 5; DB 2; Length 456;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
|||||  
Db 3 TWFLR 7

Search completed: August 31, 2006, 10:48:03  
Job time : 19.25 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:29:54 ; Search time 107.75 Seconds  
(without alignments)  
38.190 Million cell updates/sec

**Title:** DENGUE SEROTYPE2

Perfect score:

Sequence: 1 ietwflrhp 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2589679 seqs, 457216429 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2570098

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Database : A Geneseq 8:★

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

9: geneseqp2005s:\*

10: geneseqp2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score			DB	ID	Description
		Match	Length				
1	9	100.0	39	9	ADW12582	Adw12582 M1-40/DEN	
2	9	100.0	48	9	ADW12588	Adw12588 p(95-114)	
3	6	66.7	111	6	ABP75900	Abp75900 Human sec	
4	6	66.7	146	5	ADK36828	Adk36828 Novel hum	
5	6	66.7	198	6	ABR58403	AbR58403 Human NOV	
6	6	66.7	201	5	ABR90287	AbR90287 Human pol	
7	6	66.7	216	4	AAU39930	Aau39930 Human pol	
8	6	66.7	234	4	AAU29056	Aau29056 Human PRO	
9	6	66.7	234	4	AAK39929	Aam39929 Human pol	
10	6	66.7	234	4	AAK87532	Aab87532 Human PRO	
11	6	66.7	234	5	ABG95857	Abg95857 Human sec	
12	6	66.7	234	5	ABB84847	Abb84847 Human PRO	
13	6	66.7	234	5	ABB95453	Abb95453 Human ang	
14	6	66.7	234	6	ABU58432	Abu58432 Human PRO	
15	6	66.7	234	6	ABU87980	Abu87980 Human hum	
16	6	66.7	234	6	ABU84295	Abu84295 Novel hum	
17	6	66.7	234	6	ABR66169	AbR66169 Human sec	
18	6	66.7	234	6	ABR65559	AbR65559 Human sec	
19	6	66.7	234	6	ABU99499	Abu99499 Human sec	
20	6	66.7	234	6	ABU82738	Abu82738 Human PRO	
21	6	66.7	234	6	ABU89859	Abu89859 Novel hum	
22	6	66.7	234	6	ABR68108	AbR68108 Human sec	
23	6	66.7	234	6	ABU96161	Abu96161 Novel hum	

```
97 6 66.7 234 6 ABU85915 Novel hum
98 6 66.7 234 6 ABU82202 Novel hum
99 6 66.7 234 6 ABU87213 Human PRO
100 6 66.7 234 6 ABU83685 Human sec
101 6 66.7 234 6 ABO08059 Human PRO
102 6 66.7 234 6 ABU92482 Human sec
103 6 66.7 234 6 ABU81770 Novel hum
104 6 66.7 234 6 ABU65934 Novel hum
105 6 66.7 234 6 ABU81152 Human sec
106 6 66.7 234 6 ABR59763 Human sec
107 6 66.7 234 6 ABU93951 Novel hum
108 6 66.7 234 6 ABU99804 Novel hum
109 6 66.7 234 6 ABR66474 Human sec
110 6 66.7 234 6 ABR90892 Human sec
111 6 66.7 234 6 ABO53267 Novel hum
112 6 66.7 234 6 ABU943119 Human PRO
113 6 66.7 234 6 ABU79201 Human PRO
114 6 66.7 234 6 ABU86530 Human sec
115 6 66.7 234 6 ABU86835 Novel hum
116 6 66.7 234 6 ABU94624 Human PRO
117 6 66.7 234 6 ABO04551 Human PRO
118 6 66.7 234 6 ABR70300 Human sec
119 6 66.7 234 6 ABU98465 Human PRO
120 6 66.7 234 6 ABR65864 Human sec
121 6 66.7 234 6 ABR64581 Human sec
122 6 66.7 234 6 ABU79506 Human PRO
123 6 66.7 234 6 ABU92897 Human sec
124 6 66.7 234 6 ABU95856 Human PRO
125 6 66.7 234 6 ABU91076 Novel hum
126 6 66.7 234 6 ABU90169 Novel hum
127 6 66.7 234 6 ABO09584 Human sec
128 6 66.7 234 6 ABR58404 Human NOV
129 6 66.7 234 6 ABO10856 Human sec
130 6 66.7 234 6 ABR70910 Human sec
131 6 66.7 234 6 ABU98269 Novel hum
132 6 66.7 234 6 ABU87518 Human PRO
133 6 66.7 234 6 ABU91386 Human PRO
134 6 66.7 234 6 ABU89274 Novel hum
135 6 66.7 234 6 ABU84600 Human sec
136 6 66.7 234 6 ABR69690 Human sec
137 6 66.7 234 6 ABU80067 Human PRO
138 6 66.7 234 6 ABU82481 Novel hum
139 6 66.7 234 6 ABU93336 Human PRO
140 6 66.7 234 6 ABO09889 Human sec
141 6 66.7 234 6 ABO08974 Human sec
142 6 66.7 234 6 ABU96445 Human PRO
143 6 66.7 234 6 ABU10542 Human PRO
144 6 66.7 234 6 ABU72115 Human PRO
145 6 66.7 234 6 ABU95551 Human PRO
146 6 66.7 234 6 ABU96760 Novel hum
147 6 66.7 234 6 ABR70605 Human sec
148 6 66.7 234 6 ABO04956 Novel hum
149 6 66.7 234 6 ABO08364 Human sec
150 6 66.7 234 6 ABO05571 Human sec
```

## ALIGNMENTS

```
RESULT 1
ADW12582
ID ADW12582 standard; peptide; 39 AA.
XX
AC ADW12582;
XX
DT 24-MAR-2005 (first entry)
XX
DE M1-40/DEN-2 (F36) mutant protein.
XX
KW Gene therapy; protein purification; virucide; cytostatic; vaccine;
KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
KW DEN; dengue; mutant; mutein.
XX
XX
```

```
OS Dengue virus.
XX
PN US2004266987-A1.
XX
PD 30-DEC-2004.
XX
PF 30-JUN-2003; 2003US-00608029.
XX
XX
```

US2004266987-A1.

30-DEC-2004.

30-JUN-2003; 2003US-00608029.

30-JUN-2003; 2003US-00608029.

(INSP ) INST PASTEUR.

Despres P, Catteau A;

WPI; 2005-047647/05.

New isolated and purified ApoptoM peptide comprises 9 amino acids, useful as a vaccine for preventing or treating pathological conditions from non-specific febrile illnesses to severe hemorrhagic manifestations or encephalitic syndromes.

Example 1; SEQ ID NO 29; 30pp; English.

The present invention relates to an isolated and purified ApoptoM peptide. The invention is useful as a vaccine for the prevention and treatment of pathological conditions from non-specific febrile illnesses to severe hemorrhagic manifestations, encephalitic syndromes and these pathological conditions are linked to Flavivirus infection or cancers. The invention is also useful in gene therapy. The present sequence is a M1-40/DEN (dengue)-2 (F36) mutant protein.

Sequence 39 AA;

Query Match 100.0%; Score 9; DB 9; Length 39;

Best Local Similarity 100.0%; Pred. No. 0.0015;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9

Db 31 IETWFLRHP 39

|||||

RESULT 2

ADW12588

ID ADW12588 standard; protein; 48 AA.

XX

AC ADW12588;

XX

DT 24-MAR-2005 (first entry)

XX

DE p(95-114) EGFP(M1-M40)DEN-2 (I36F) plasmid DNA encoded protein #3.

XX

KW Gene therapy; protein purification; virucide; cytostatic; vaccine; hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer; DEN; dengue; EGFP; enhanced green fluorescent protein.

XX

OS Dengue virus.

OS Chimeric.

OS Unidentified.

XX

Key Location/Qualifiers

FT Misc-difference 2 /note= "Encoded by GGC"

FT Misc-difference 4 /note= "Encoded by GAC"

FT Misc-difference 13.44 /note= "Encoded by GTTTC"

XX

PN US2004266987-A1.

XX

PD 30-DEC-2004.

XX

PF 30-JUN-2003; 2003US-00608029.



```

XX PR 30-JUN-2003; 2003US-00608029.
XX PA (INSP ) INST PASTEUR.
XX PI
XX PI Despres P, Catteau A;
XX DR WPI; 2005-047647/05.
XX DR N-PSDB; ADW12589.
XX PT New isolated and purified Apoptom peptide comprises 9 amino acids, useful
XX PT as a vaccine for preventing or treating pathological conditions from non-
XX PT specific febrile illnesses to severe hemorrhagic manifestations or
XX PT encephalitic syndromes.
XX PS Disclosure; SEQ ID NO 35; 30pp; English.
XX CC The present invention relates to an isolated and purified Apoptom
XX CC peptide. The invention is useful as a vaccine for the prevention and
XX CC treatment of pathological conditions from non-specific febrile illnesses
XX CC to severe hemorrhagic manifestations, encephalitic syndromes and these
XX CC pathological conditions are linked to Flavivirus infection or cancers.
XX CC The invention is also useful in gene therapy. The present sequence is a
XX CC p(95-114) EGFP (enhanced green fluorescent protein) (M1-M40)DEN (dengue)-2
XX CC (136F) plasmid DNA encoded protein.
XX SQ Sequence 48 AA;
XX
XX Query Match 100.0%; Score 9; DB 9; Length 48;
XX Best Local Similarity 100.0%; Pred. No. 0.0019;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 IETWFLRHP 9
XX | | | | |
XX 40 IETWFLRHP 48
XX
XX DB
XX
XX RESULT 3
XX ABP75900
XX ID ABP75900 standard; protein; 111 AA.
XX AC ABP75900;
XX DT 10-FEB-2003 (first entry)
XX DE Human secretory polypeptide SPTM SEQ ID NO 1084.
XX
XX KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
XX KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
XX KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
XX KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
XX KW anti-inflammatory; immunosuppressive; neuroprotective; nontropic;
XX KW neuroleptic; anticonvulsant; cytotatic; antiparkinsonian; anxiolytic;
XX KW antipsoriatic; antianemic; anti-HIV; human immunodeficiency virus;
XX KW secretory polynucleotide; secretory protein.
XX OS Homo sapiens.
XX
XX PN WO200283876-A2.
XX PD 24-OCT-2002.
XX
XX PF 27-MAR-2002; 2002WO-US009921.
XX
XX PR 29-MAR-2001; 2001US-0280067P.
XX PR 29-MAR-2001; 2001US-0280068P.
XX PR 16-MAY-2001; 2001US-0291280P.
XX PR 17-MAY-2001; 2001US-0291829P.
XX PR 17-MAY-2001; 2001US-0291849P.
XX PR 19-JUN-2001; 2001US-0299428P.
XX PR 20-JUN-2001; 2001US-0299776P.
XX PR 20-JUN-2001; 2001US-0300001P.
XX
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX PI Dufour GE, Hillman JB, Yu JY, Tuason O, Yap PE, Anshey SR;
XX PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
XX PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-075543/07.
XX DR N-PSDB; ABZ36342.
XX
XX PT New human secretory proteins and polynucleotides, useful for diagnosing,
XX PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
XX PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
XX PT cancers.
XX PS Claim 27; SEQ ID NO 1084; 458pp + Sequence Listing; English.
XX
XX CC The invention relates to a secretory polynucleotide (designated sptm)
XX CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
XX CC naturally occurring polynucleotide sequence at least 90 % identical to
XX CC the polynucleotide sequence, a polynucleotide complementary to them or an
XX CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
XX CC treating, preventing or diagnosing a disease or condition associated with
XX CC the expression of functional SPTM. These are particularly useful for
XX CC diagnosing, treating or preventing autoimmune/inflammatory disorders
XX CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
XX CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
XX CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
XX CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
XX CC schizophrenia or amnesia), or cell proliferative disorders (e.g.
XX CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,
XX CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
XX CC breast, cervix or prostate). The present sequence is one of the SPTM
XX CC proteins of the invention (ABP75384-ABP75962). Note: The sequence data
XX CC for this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 111 AA;
XX
XX Query Match 66.7%; Score 6; DB 6; Length 111;
XX Best Local Similarity 100.0%; Pred. No. 14;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 IETWFL 6
XX | | | | |
XX 43 IETWFL 48
XX
XX DB
XX
XX RESULT 4
XX ADK36828
XX ID ADK36828 standard; protein; 146 AA.
XX AC ADK36828;
XX DT 06-MAY-2004 (first entry)
XX DE Novel human polypeptide SeqID8910.
XX
XX KW antiarthritic; antiparkinsonian; neuroprotective; nontropic;
XX KW immunosuppressive; cytotatic; antipsoriatic; antiinflammatory;
XX KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
XX KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
XX KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
XX KW fungus; parasite; human.
XX OS Homo sapiens.
XX
XX PH Key Location/Qualifiers
XX FT Misc-difference 1..146
XX FT /label= OTHER
XX FT /note= "OTHER= All Xaa's in this sequence are unknown
XX FT

```

FT amino acids or the site of a stop codon within the DNA  
FT sequence"  
XX WO200216439-A2.  
PN  
XX  
XX 28-FEB-2002.  
PD  
XX  
XX 05-MAR-2001; 2001WO-US004941.  
PF  
XX  
XX 07-MAR-2000; 2000US-00519705.  
PR  
XX 19-MAY-2000; 2000US-00574454.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Tang YT, Liu C, Drmanac RT;  
PI WPI; 2002-280918/32.  
XX  
XX Isolated polynucleotide encoding bone marrow derived polypeptides useful  
PT for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's  
PT disease, and inflammatory bowel disease.  
XX  
XX Claim 20; SEQ ID NO 8910; 504pp; English.  
PS  
XX  
XX This invention relates to a novel isolated polynucleotide comprising a  
CC nucleotide sequence selected from one of 1680 sequences, a mature protein  
CC coding portion of them, an active domain of them and their complementary  
CC sequences. The invention may be useful for the production of compounds  
CC with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,  
CC immunosuppressive, cytostatic, antipsoriatic, antiinflammatory,  
CC antibacterial, antiviral, antifungal or antiparasitic activity. In  
CC addition, the disclosed sequences may be useful for gene therapy. The  
CC polypeptides or their antibodies are useful for treating many diseases  
CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,  
CC psoriasis, inflammatory bowel disease and infections caused by bacteria,  
CC viruses, fungi or parasites. The present sequence is that of a human  
CC polypeptide of the invention.  
XX  
XX Sequence 146 AA;  
SQ  
Query Match 66.7%; Score 6; DB 5; Length 146;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IETWFL 6  
Db 116 IETWFL 121  
RESULT 5  
ABR58403  
ID ABR58403 standard; protein; 198 AA.  
XX  
XX ABR58403;  
AC  
XX  
XX 07-JUL-2003 (first entry)  
DT  
XX  
XX Human NOV19a.  
DE  
XX  
XX Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;  
KW antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;  
KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;  
KW neurodegenerative disorder; Alzheimer's disease; immune disorder;  
KW haematopoietic disorder.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2003029423-A2.  
PN  
XX  
XX 10-APR-2003.  
PD  
XX  
XX 02-OCT-2002; 2002WO-US031358.  
PF

XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327342P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 12-APR-2002; 2002US-0371972P.  
PR 12-APR-2002; 2002US-0371980P.  
PR 17-APR-2002; 2002US-0373261P.  
PR 19-APR-2002; 2002US-0373805P.  
PR 23-APR-2002; 2002US-0374738P.  
PR 16-MAY-2002; 2002US-0381101P.  
PR 17-MAY-2002; 2002US-0381635P.  
PR 29-MAY-2002; 2002US-0383830P.  
PR 01-OCT-2002; 2002US-00262839.  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX  
XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;  
PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;  
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;  
XX  
XX WPI; 2003-381625/36.  
DR N-PSDB; ACC72115.  
DR  
XX  
XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
PT dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
PT  
XX  
XX Claim 1; Page 165; 487pp; English.  
PS  
XX  
XX The present invention relates to novel human NOV proteins and their  
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
CC proteins are useful in manufacturing a medicament for treating a syndrome  
CC associated with a human disease. The NOV proteins and coding sequences  
CC may be used to diagnose, treat or prevent metabolic disorders such as  
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
CC disorders, haematopoietic disorders and various dyslipidaemias  
XX  
XX Sequence 198 AA;  
SQ  
Query Match 66.7%; Score 6; DB 6; Length 198;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IETWFL 6  
Db 130 IETWFL 135  
RESULT 6  
ABB90287  
ID ABB90287 standard; protein; 201 AA.  
XX  
XX ABB90287;  
AC  
XX  
XX 24-MAY-2002 (first entry)  
DT  
XX  
XX Human polypeptide SEQ ID NO 2663.  
DE  
XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
 KW antineurotic; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 XX Homo sapiens.  
 OS  
 PN WO200190304-A2.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001WO-US016450.  
 XX  
 PR 19-MAY-2000; 2000US-0205515P.  
 XX  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Birse CE, Rosen CA;  
 XX WPI; 2002-122018/16.  
 DR N-PSDB; ABL90696.  
 XX  
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
 PT prevention of neural, immune system, muscular, reproductive,  
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
 PT disorders.  
 XX  
 PS Claim 11; SEQ ID NO 2663; 2081pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (ABL9449-ABL90853) and proteins  
 CC (ABB9040-ABB90444) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 201 AA;  
 Query Match 66.7%; Score 6; DB 5; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IETWFL 6  
 Db 133 IETWFL 138  
 RESULT 7  
 AAM39930  
 ID AAM39930 standard; protein; 216 AA.  
 XX  
 AC AAM39930;  
 XX  
 XX 22-OCT-2001 (first entry)  
 DT  
 XX Human polypeptide SEQ ID NO 3075.  
 DE  
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX (HYSE-) HYSEQ INC.  
 PA  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX WPI; 2001-442253/47.  
 DR N-PSDB; AAI59086.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX  
 PS Example 4; SEQ ID NO 3075; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX  
 SQ Sequence 216 AA;  
 Query Match 66.7%; Score 6; DB 4; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IETWFL 6  
 Db 148 IETWFL 153  
 RESULT 8  
 AAU29056  
 ID AAU29056 standard; protein; 234 AA.  
 XX  
 AC AAU29056;  
 XX  
 XX 18-DEC-2001 (first entry)  
 DT  
 XX Human PRO polypeptide sequence #33.  
 DE  
 XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;

KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
 XX

OS Homo sapiens.

PN WO200168848-A2.

XX 20-SEP-2001.

XX 28-FEB-2001; 2001WO-US006520.

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005841.

PR 03-MAR-2000; 2000US-0187202P.

PR 06-MAR-2000; 2000US-0186968P.

PR 14-MAR-2000; 2000US-0189320P.

PR 14-MAR-2000; 2000US-0189328P.

PR 15-MAR-2000; 2000WO-US006884.

PR 21-MAR-2000; 2000US-0190828P.

PR 21-MAR-2000; 2000US-0191007P.

PR 21-MAR-2000; 2000US-0191048P.

PR 21-MAR-2000; 2000US-0191314P.

PR 28-MAR-2000; 2000US-0192655P.

PR 29-MAR-2000; 2000US-0193032P.

PR 29-MAR-2000; 2000US-0193053P.

PR 30-MAR-2000; 2000WO-US008439.

PR 04-APR-2000; 2000US-0191048P.

PR 04-APR-2000; 2000US-0194449P.

PR 11-APR-2000; 2000US-0195975P.

PR 11-APR-2000; 2000US-0196000P.

PR 11-APR-2000; 2000US-0196187P.

PR 11-APR-2000; 2000US-0196590P.

PR 11-APR-2000; 2000US-0196820P.

PR 18-APR-2000; 2000US-0198121P.

PR 18-APR-2000; 2000US-0198585P.

PR 25-APR-2000; 2000US-0199397P.

PR 25-APR-2000; 2000US-0199550P.

PR 25-APR-2000; 2000US-0199654P.

PR 03-MAY-2000; 2000US-0201516P.

PR 17-MAY-2000; 2000WO-US013705.

PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.

PR 05-JUN-2000; 2000US-0209832P.

PR 28-JUL-2000; 2000WO-US020710.

PR 22-AUG-2000; 2000US-00644848.

CC and rabbits but are preferably human. The polypeptides can be used to  
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood,  
 CC when contacted with it. A specific polypeptide can be used to stimulate  
 CC the proliferation or differentiation of chondrocyte cells. The PRO  
 CC proteins can be used to determine the presence of tumours and also  
 CC susceptibility to tumour development, particularly adrenal, lung, colon,  
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian  
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids  
 CC can be used for genetic analysis of individuals with genetic disorders  
 XX  
 XX Sequence 234 AA;

Query Match 66.7%; Score 6; DB 4; Length 234;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6

Db 166 IETWFL 171

RESULT 9

AAM39929

ID AAM39929 standard; protein; 234 AA.

AC AAM39929;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 3074.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX Homo sapiens.

OS WO200153312-A1.

PN 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

PI Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

DR N-PSDB; AAI59085.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

CC The PRO polypeptides and their associated nucleic acids can be used to  
 CC detect the presence of a tumour in a mammal by comparing the level of  
 CC expression of a PRO polypeptide in a test sample of cells from the animal  
 CC and a control sample of normal cells, whereby a higher level of  
 CC expression in the test sample indicates the presence of a tumour in the  
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats

XX Claim 11; Fig 66; 774pp; English.

XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.

CC The PRO polypeptides and their associated nucleic acids can be used to

CC detect the presence of a tumour in a mammal by comparing the level of

CC expression of a PRO polypeptide in a test sample of cells from the animal

CC and a control sample of normal cells, whereby a higher level of

CC expression in the test sample indicates the presence of a tumour in the

CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the

PT presence of tumors, such as prostate and breast tumors, in mammals and to

PT screen for modulators of the compounds.

XX Claim 11; Fig 66; 774pp; English.

PS Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.

XX The PRO polypeptides and their associated nucleic acids can be used to

CC detect the presence of a tumour in a mammal by comparing the level of

CC expression of a PRO polypeptide in a test sample of cells from the animal

CC and a control sample of normal cells, whereby a higher level of

CC expression in the test sample indicates the presence of a tumour in the

CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the

PT presence of tumors, such as prostate and breast tumors, in mammals and to

PT screen for modulators of the compounds.

XX Claim 11; Fig 66; 774pp; English.

PS Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.

XX The PRO polypeptides and their associated nucleic acids can be used to

CC detect the presence of a tumour in a mammal by comparing the level of

CC expression of a PRO polypeptide in a test sample of cells from the animal

CC and a control sample of normal cells, whereby a higher level of

CC expression in the test sample indicates the presence of a tumour in the

CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the

PT presence of tumors, such as prostate and breast tumors, in mammals and to

PT screen for modulators of the compounds.

XX Claim 11; Fig 66; 774pp; English.

PS Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.

XX The PRO polypeptides and their associated nucleic acids can be used to

CC detect the presence of a tumour in a mammal by comparing the level of

CC expression of a PRO polypeptide in a test sample of cells from the animal

CC and a control sample of normal cells, whereby a higher level of

CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification

XX SQ Sequence 234 AA;

Query Match 66.7%; Score 6; DB 4; Length 234;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWFL 6  
 |||||  
 Db 166 IETWFL 171

RESULT 10  
 AAB87532  
 ID AAB87532 standard; protein; 234 AA.

XX AC AAB87532;  
 XX DT 15-MAY-2001 (first entry)  
 XX DE Human PRO1864.  
 XX DE Human; PRO protein; mapping.

XX OS Homo sapiens.

XX PN WO200116318-A2.

XX PD 08-MAR-2001.

XX PF 24-AUG-2000; 2000WO-US023328.  
 XX PR 01-SEP-1999; 99WO-US020111.  
 XX PR 15-SEP-1999; 99WO-US021090.  
 XX PR 07-DEC-1999; 99US-0169495P.  
 XX PR 09-DEC-1999; 99US-0170262P.  
 XX PR 11-JAN-2000; 2000US-0175481P.  
 XX PR 18-FEB-2000; 2000WO-US0004341.  
 XX PR 18-FEB-2000; 2000WO-US0004342.  
 XX PR 22-FEB-2000; 2000WO-US0004414.  
 XX PR 01-MAR-2000; 2000WO-US0005601.  
 XX PR 03-MAR-2000; 2000US-0187202P.  
 XX PR 21-MAR-2000; 2000US-0191007P.  
 XX PR 30-MAR-2000; 2000WO-US0008439.  
 XX PR 25-APR-2000; 2000US-0199397P.  
 XX PR 22-MAY-2000; 2000WO-US014042.  
 XX PR 05-JUN-2000; 2000US-0209832P.

XX PA (GETH ) GENENTECH INC.

XX PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski RJ;  
 XX PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;

XX DR WPI; 2001-183260/18.  
 XX DR N-PSDB; AAF92064.

XX PT Eighty four nucleic acids encoding PRO polypeptides, useful in molecular  
 XX PT biology, including use as hybridization probes, and in chromosome and  
 XX PT gene mapping.

XX PS Claim 12; Fig 14; 278pp; English.

XX CC The present sequence is a human PRO polypeptide (secreted and  
 CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or  
 CC anti-PRO antibodies are useful for preparation of a medicament useful in  
 CC the treatment of a condition which is responsive to the PRO protein,  
 CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be  
 CC employed as molecular weight markers for protein electrophoresis. The PRO  
 CC coding sequence has applications in molecular biology, including use as  
 CC hybridisation probes, and in chromosome and gene mapping

XX SQ Sequence 234 AA;

Query Match 66.7%; Score 6; DB 4; Length 234;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWFL 6  
 |||||  
 Db 166 IETWFL 171

RESULT 11  
 ABG95857  
 ID ABG95857 standard; protein; 234 AA.

XX AC ABG95857;  
 XX DT 10-DEC-2002 (first entry)

XX DE Human secreted/transmembrane protein PRO1864.

XX KW Human; secreted protein; transmembrane protein; antirheumatic;  
 XX KW antiarthritic; osteopathic; sports-related joint problem;  
 XX KW articular cartilage defect; osteoarthritis; rheumatoid arthritis.

XX OS Homo sapiens.

XX PN US2002119130-A1.

XX PD 29-AUG-2002.

XX PF 06-DEC-2001; 2001US-0006867.

XX PR 29-OCT-1997; 97US-0063435P.  
 XX PR 29-OCT-1997; 97US-0064215P.  
 XX PR 22-APR-1998; 98US-0082797P.  
 XX PR 29-APR-1998; 98US-0083495P.  
 XX PR 15-MAY-1998; 98US-0085579P.  
 XX PR 02-JUN-1998; 98US-0087759P.  
 XX PR 04-JUN-1998; 98US-0088021P.  
 XX PR 04-JUN-1998; 98US-0088029P.  
 XX PR 04-JUN-1998; 98US-0088030P.  
 XX PR 10-JUN-1998; 98US-0088734P.  
 XX PR 10-JUN-1998; 98US-0088740P.  
 XX PR 10-JUN-1998; 98US-0088811P.  
 XX PR 10-JUN-1998; 98US-0088824P.  
 XX PR 10-JUN-1998; 98US-0088825P.  
 XX PR 11-JUN-1998; 98US-0088863P.  
 XX PR 12-JUN-1998; 98US-0089105P.  
 XX PR 16-JUN-1998; 98US-0089514P.  
 XX PR 17-JUN-1998; 98US-0089653P.  
 XX PR 22-JUN-1998; 98US-0090246P.  
 XX PR 19-JUN-1998; 98US-0089952P.  
 XX PR 24-JUN-1998; 98US-0090444P.  
 XX PR 25-JUN-1998; 98US-0090688P.  
 XX PR 25-JUN-1998; 98US-0090696P.  
 XX PR 26-JUN-1998; 98US-0090862P.  
 XX PR 02-JUL-1998; 98US-0091628P.  
 XX PR 10-AUG-1998; 98US-0096012P.  
 XX PR 17-AUG-1998; 98US-0096757P.  
 XX PR 18-AUG-1998; 98US-0096949P.  
 XX PR 18-AUG-1998; 98US-0096959P.  
 XX PR 26-AUG-1998; 98US-0097954P.

PR 26-AUG-1998; 98US-0097971P.  
 PR 26-AUG-1998; 98US-0097979P.  
 PR 01-SEP-1998; 98US-0098749P.  
 PR 10-SEP-1998; 98US-0099741P.  
 PR 10-SEP-1998; 98US-0099763P.  
 PR 10-SEP-1998; 98US-0099792P.  
 PR 10-SEP-1998; 98US-0099812P.  
 PR 10-SEP-1998; 98US-0099815P.  
 PR 16-SEP-1998; 98US-0100627P.  
 PR 16-SEP-1998; 98US-0100662P.  
 PR 16-SEP-1998; 98US-0101933P.  
 PR 17-SEP-1998; 98US-0100683P.  
 PR 17-SEP-1998; 98US-0100684P.  
 PR 17-SEP-1998; 98US-0100930P.  
 PR 22-SEP-1998; 98US-0101279P.  
 PR 23-SEP-1998; 98US-0101475P.  
 PR 24-SEP-1998; 98US-0101738P.  
 PR 24-SEP-1998; 98US-0101743P.  
 PR 24-SEP-1998; 98US-0101916P.  
 PR 20-SEP-1998; 98US-0102570P.  
 PR 06-OCT-1998; 98US-0103449P.  
 PR 14-MAY-1999; 99WO-US005028.  
 PR 18-MAY-1999; 99WO-US010733.  
 PR 02-JUN-1999; 99WO-US012252.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 15-SEP-1999; 99WO-US021194.  
 PR 22-DEC-1999; 99WO-US030720.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 01-DEC-2000; 2000WO-US032378.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-MAR-2001; 2001WO-US006666.  
 PR 30-MAY-2001; 2001WO-US017443.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 XX (GETH ) GENENTECH INC.  
 PA Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;  
 XX WPI; 2002-731348/79.  
 DR N-PSDB; ABS74384.  
 DR  
 XX  
 XX  
 XX  
 PT New isolated secreted and transmembrane PRO polypeptide useful for  
 PT modulating biological activity of a cell, or for treating sports-related  
 PT joint problems, osteoarthritis or rheumatoid arthritis.  
 XX  
 XX  
 PS Claim 20; Fig 14; 399pp; English.  
 XX  
 XX  
 CC The invention relates to an isolated secreted and transmembrane PRO  
 CC polypeptide having 80 % sequence identity to a sequence appearing as  
 CC ABG5851-ABG5934 or their associated signal peptide, or a sequence of an  
 CC extracellular domain of the proteins with their associated signal peptide  
 CC or lacking its associated signal peptide. Also included are the nucleic  
 CC acids encoding the proteins, vectors, host cells, fusion proteins and  
 CC antibodies which specifically bind to the proteins. The proteins are  
 CC useful for detecting a polypeptide designated as A, B, C or D in a sample  
 CC suspected of containing A, B, C or D polypeptide, by contacting the  
 CC sample with a polypeptide designated as E, F, G, H or I (or vice versa)  
 CC and determining the formation of a A/E, B/F, B/G, C/H or D/I polypeptide

CC conjugate in the sample, where the formation of the conjugate is  
 CC indicative of the presence of an A, B, C or D polypeptide in the sample,  
 CC where A is a PRO10272 polypeptide, B is a PRO20110 polypeptide, C is a  
 CC PRO10096 polypeptide, D is a PRO19760 polypeptide, E is a PRO5801  
 CC polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a  
 CC PRO20233 polypeptide and I is a PRO1890 polypeptide. The sample comprises  
 CC a cell suspected of expressing the A, B, C or D polypeptide. The E, F, G,  
 CC H or I polypeptide is labeled with a detectable label or is attached to a  
 CC solid support. The proteins are useful for linking a bioactive molecule  
 CC to a cell expressing a polypeptide designated as A, B, C or D or E, F, G,  
 CC H or I. The bioactive molecule is a toxin, a radiolabel or an antibody.  
 CC The bioactive molecule causes death of the cell. A, B, C, D, E, F, G, H,  
 CC or I, or antibodies against them are useful for modulating a biological  
 CC activity of a cell expressing a polypeptide designated as A, B, C or D or  
 CC E, F, G, H, or I. The cell is killed. The proteins are useful for  
 CC identifying agonists or antagonists, for the preparation of a medicament  
 CC useful in the treatment of a condition which is responsive to the  
 CC proteins, as molecular weight markers for protein electrophoresis  
 CC purposes, and as therapeutic agents for treating sports-related joint  
 CC problems, articular cartilage defects, osteoarthritis or rheumatoid  
 CC arthritis. Nucleic acids encoding the proteins are useful as  
 CC hybridisation probes, in chromosome and gene mapping, in the generation  
 CC of anti-sense RNA and DNA, for the preparation of the proteins, to  
 CC generate transgenic or knockout animals which are useful in the  
 CC development and screening of therapeutic useful reagents, for chromosome  
 CC identification, and in gene therapy. The antibody is useful as a  
 CC therapeutic agent, in a diagnostic assay and for affinity purification of  
 CC the protein from recombinant cell culture natural sources. The present  
 CC invention represents a novel secreted or transmembrane protein of the

Query Match 66.7%; Score 6; DB 5; Length 234;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6

Db 166 IETWFL 171

RESULT 12

ABB84847

ID ABB84847 standard; protein; 234 AA.

XX ABB84847;

XX

DT 16-MAY-2002 (first entry)

DE Human PRO1864 protein sequence SEQ ID NO:62.

XX

KW Human; angiogenesis; cardiant; cytostatic; antiangiogenic; hypotensive;  
 KW vulnary; antiarteriosclerotic; PRO agonist; PRO antagonist; trauma;  
 KW gene therapy; cardiovascular disorder; endothelial disorder; cancer;  
 KW angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;  
 KW age-related macular degeneration; arterial restenosis; angina;  
 KW rheumatoid arthritis; myocardial infarction; thrombophlebitis;  
 KW lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;  
 KW wound healing; chromosome mapping; gene mapping.

OS Homo sapiens.

XX

PN WO200200690-A2.

XX

PD 03-JAN-2002.

XX

PF 20-JUN-2001; 2001WO-US019692.

XX

PR 23-JUN-2000; 2000US-0213637P.

PR

PR 20-JUL-2000; 2000US-0219556P.

PR

PR 25-JUL-2000; 2000US-0220624P.

PR

PR 25-JUL-2000; 2000US-0220664P.

PR 28-JUL-2000; 2000WO-US020710.  
PR 02-AUG-2000; 2000US-0222695P.  
PR 17-AUG-2000; 2000US-00643657.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 07-SEP-2000; 2000US-0230978P.  
PR 18-SEP-2000; 2000US-00664610.  
PR 18-SEP-2000; 2000US-00665350.  
PR 24-OCT-2000; 2000US-0242922P.  
PR 08-NOV-2000; 2000US-00709238.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00643657.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 22-JAN-2001; 2001US-00767609.  
PR 28-FEB-2001; 2001US-00796498.  
PR 01-MAR-2001; 2001WO-US006520.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 30-MAY-2001; 2001US-00870574.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
XX (GETH ) GENENTECH INC.  
PA  
XX Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;  
PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;  
XX WPI; 2002-090516/12.  
DR N-PSDB; ABL88102.  
XX  
PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial  
PT infarction), endothelial or angiogenic disorders in a mammal.  
XX  
PS Claim 11; Fig 62; 565pp; English.  
XX  
CC ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to  
CC ABB85003. The PRO proteins and polynucleotides have cardiant, cytostatic,  
CC antiangiogenic, hypotensive, vulnerary and antiarteriosclerotic  
CC activities, and can be used in gene therapy. The PRO polynucleotides,  
CC proteins, agonists and antagonists are useful for treating or diagnosing  
CC a cardiovascular, endothelial or angiogenic disorder in a mammal, e.g.  
CC cardiac hypertrophy, trauma, cancer, age-related macular degeneration,  
CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,  
CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour  
CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound  
CC healing. The PRO polynucleotides have applications in molecular biology,  
CC including use as hybridisation probes, and in chromosome and gene  
CC mapping. ABL88259 to ABL88267 represent primers and probes used in the  
CC exemplification of the present invention  
XX  
SQ Sequence 234 AA;  
  
Query Match 66.7%; Score 6; DB 5; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 IETWFL 6  
DB 166 IETWFL 171

RESULT 13  
ABB95453  
ID ABB95453 standard; protein; 234 AA.  
XX  
AC ABB95453;  
XX  
DT 19-JUL-2002 (first entry)  
XX  
DE Human angiogenesis related protein PRO1864 SEQ ID NO: 62.  
XX  
KW Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;  
KW atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;  
KW cardiant; cytostatic; antiangiogenic; hypotensive; vulnerary;  
KW antiarteriosclerotic.  
XX  
OS Homo sapiens.  
XX  
PN WO200208284-A2.  
XX  
PD 31-JAN-2002.  
XX  
PF 09-JUL-2001; 2001WO-US021735.  
XX  
PR 20-JUL-2000; 2000US-0219556P.  
PR 25-JUL-2000; 2000US-0220624P.  
PR 25-JUL-2000; 2000US-0220664P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 02-AUG-2000; 2000US-0222695P.  
PR 17-AUG-2000; 2000US-00643657.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 07-SEP-2000; 2000US-0230978P.  
PR 18-SEP-2000; 2000US-00664610.  
PR 18-SEP-2000; 2000US-00665350.  
PR 24-OCT-2000; 2000US-0242922P.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 22-JAN-2001; 2001US-00767609.  
PR 28-FEB-2001; 2001US-00796498.  
PR 01-MAR-2001; 2001WO-US006520.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 30-MAY-2001; 2001US-00870574.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
XX (GETH ) GENENTECH INC.  
PA (BAKE/) BAKER K P.  
PA (FERR/) FERRARA N.  
PA (GERB/) GERBER H.  
PA (GERR/) GERRITSEN M E.  
PA (GODD/) GODDARD A.  
PA (GODO/) GOWSKI P J.  
PA (GURN/) GURNEY A L.  
PA (HILL/) HILLAN K J.  
PA (MARS/) MARSTERS S A.  
PA (PANJ/) PAN J.  
PA (PAONI/) PAONI N F.  
PA (STEP/) STEPHAN J F.  
PA (WATA/) WATANABE C K.

PA (WILL/) WILLIAMS P M.  
PA (WOOD/) WOOD W I.  
XX Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;  
PI Stephens JF, Watanabe CK, Williams PM, Wood WT, Ye W;  
XX WPI: 2002-171999/22.  
DR N-PSDB; ABL95591.  
XX  
PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial  
PT infarction), endothelial or angiogenic disorders in a mammal.  
XX  
PS Claim 11; Fig 62; 567pp; English.  
XX  
CC The present invention provides the protein and coding sequences of human  
CC PRO proteins. These are useful for treating or diagnosing a  
CC cardiovascular, endothelial or angiogenic disorder, including cardiac  
CC hypertrophy, trauma, cancer, age-related macular degeneration,  
CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,  
CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour  
CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound  
CC healing. The present sequence is a PRO protein of the invention  
XX  
SQ Sequence 234 AA;  
  
Query Match 66.7%; Score 6; DB 5; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 IETWFL 6  
Db 166 IETWFL 171  
  
RESULT 14  
ABU58432  
ID ABU58432 standard; protein; 234 AA.  
XX AC ABU58432;  
XX DT 15-APR-2003 (first entry)  
XX DE Human PRO polypeptide #33.  
XX  
KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;  
KW dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;  
KW antibody-dependent enzyme mediated prodrug therapy.  
XX  
XX OS Homo sapiens.  
XX  
XX US2003027272-A1.  
XX  
XX PD 06-FEB-2003.  
XX  
XX PF 21-JUN-2002; 2002US-00176492.  
XX  
XX 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
  
PR 24-NOV-1997; 97US-0066456P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0089335P.  
PR 12-DEC-1997; 97US-0089425P.  
PR 17-DEC-1997; 97US-0089870P.  
PR 18-DEC-1997; 97US-00868017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 28-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088328P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088828P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.



PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 24-JUN-1998; 98US-0090354P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090896P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091159P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 04-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-01019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100949P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 23-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.

PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-00168978.

Query Match 66.7%; Score 6; DB 6; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IETWFL 6  
|||||  
Db 166 IETWFL 171

RESULT 15  
ABU87980

ID ABU87980 standard; protein; 234 AA.

XX AC ABU87980;

XX DT 07-JUL-2003 (first entry)

XX DE Novel human secreted and transmembrane protein PRO1864.  
XX KW Human; secreted and transmembrane protein; PRO; gene therapy;  
KW tumour necrosis factor-alpha release; TNF-alpha release;  
KW chondrocyte proliferation; chondrocyte differentiation; tumour;  
KW adrenal tumour; lung tumour; colon tumour; breast tumour;  
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.

XX OS Homo sapiens.

XX PN US2003032127-A1.

XX PD 13-FEB-2003.

XX PF 26-JUN-2002; 2002US-00183012.

XX PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.

PR 18-DEC-1997; 97US-0068017P.  
 PR 10-MAR-1998; 98US-0077450P.  
 PR 11-MAR-1998; 98US-0077632P.  
 PR 11-MAR-1998; 98US-0077642P.  
 PR 20-MAR-1998; 98US-0078886P.  
 PR 20-MAR-1998; 98US-0078933P.  
 PR 27-MAR-1998; 98US-0079664P.  
 PR 27-MAR-1998; 98US-0079786P.  
 PR 31-MAR-1998; 98US-0080107P.  
 PR 31-MAR-1998; 98US-0080194P.  
 PR 01-APR-1998; 98US-0080327P.  
 PR 01-APR-1998; 98US-0080333P.  
 PR 08-APR-1998; 98US-0081049P.  
 PR 08-APR-1998; 98US-0081070P.  
 PR 09-APR-1998; 98US-0081195P.  
 PR 15-APR-1998; 98US-0081838P.  
 PR 21-APR-1998; 98US-0082568P.  
 PR 21-APR-1998; 98US-0082569P.  
 PR 22-APR-1998; 98US-0082704P.  
 PR 22-APR-1998; 98US-0082797P.  
 PR 28-APR-1998; 98US-0083322P.  
 PR 29-APR-1998; 98US-0083495P.  
 PR 29-APR-1998; 98US-0083496P.  
 PR 29-APR-1998; 98US-0083499P.  
 PR 05-MAY-1998; 98US-0084366P.  
 PR 06-MAY-1998; 98US-0084414P.  
 PR 07-MAY-1998; 98US-0084639P.  
 PR 07-MAY-1998; 98US-0084640P.  
 PR 07-MAY-1998; 98US-0084643P.  
 PR 15-MAY-1998; 98US-0085579P.  
 PR 15-MAY-1998; 98US-0085580P.  
 PR 15-MAY-1998; 98US-0085582P.  
 PR 15-MAY-1998; 98US-0085700P.  
 PR 18-MAY-1998; 98US-0086023P.  
 PR 22-MAY-1998; 98US-0086392P.  
 PR 22-MAY-1998; 98US-0086486P.  
 PR 28-MAY-1998; 98US-0087098P.  
 PR 28-MAY-1998; 98US-0087208P.  
 PR 02-JUN-1998; 98US-0087609P.  
 PR 02-JUN-1998; 98US-0087759P.  
 PR 03-JUN-1998; 98US-0087827P.  
 PR 04-JUN-1998; 98US-0088025P.  
 PR 04-JUN-1998; 98US-0088028P.  
 PR 04-JUN-1998; 98US-0088029P.  
 PR 04-JUN-1998; 98US-0088033P.  
 PR 04-JUN-1998; 98US-0088326P.  
 PR 05-JUN-1998; 98US-0088167P.  
 PR 05-JUN-1998; 98US-0088202P.  
 PR 05-JUN-1998; 98US-0088212P.  
 PR 05-JUN-1998; 98US-0088217P.  
 PR 09-JUN-1998; 98US-0088655P.  
 PR 10-JUN-1998; 98US-0088722P.  
 PR 10-JUN-1998; 98US-0088738P.  
 PR 10-JUN-1998; 98US-0088740P.  
 PR 10-JUN-1998; 98US-0088811P.  
 PR 10-JUN-1998; 98US-0088824P.  
 PR 10-JUN-1998; 98US-0088825P.  
 PR 10-JUN-1998; 98US-0088826P.  
 PR 11-JUN-1998; 98US-0088861P.  
 PR 11-JUN-1998; 98US-0088863P.  
 PR 11-JUN-1998; 98US-0088976P.  
 PR 12-JUN-1998; 98US-0089090P.  
 PR 12-JUN-1998; 98US-0089105P.  
 PR 16-JUN-1998; 98US-0089512P.  
 PR 16-JUN-1998; 98US-0089514P.  
 PR 17-JUN-1998; 98US-0089538P.  
 PR 17-JUN-1998; 98US-0089598P.  
 PR 17-JUN-1998; 98US-0089553P.  
 PR 18-JUN-1998; 98US-0089908P.  
 PR 19-JUN-1998; 98US-0089952P.  
 PR 22-JUN-1998; 98US-0090246P.  
 PR 22-JUN-1998; 98US-0090252P.

PR 22-JUN-1998; 98US-0090254P.  
 PR 24-JUN-1998; 98US-0090423P.  
 PR 24-JUN-1998; 98US-0090435P.  
 PR 24-JUN-1998; 98US-0090444P.  
 PR 24-JUN-1998; 98US-0090461P.  
 PR 24-JUN-1998; 98US-0090535P.  
 PR 24-JUN-1998; 98US-0090540P.  
 PR 25-JUN-1998; 98US-0090676P.  
 PR 25-JUN-1998; 98US-0090678P.  
 PR 25-JUN-1998; 98US-0090688P.  
 PR 25-JUN-1998; 98US-0090690P.  
 PR 25-JUN-1998; 98US-0090694P.  
 PR 25-JUN-1998; 98US-0090695P.  
 PR 25-JUN-1998; 98US-0090696P.  
 PR 26-JUN-1998; 98US-00105413.  
 PR 26-JUN-1998; 98US-0090862P.  
 PR 26-JUN-1998; 98US-0090863P.  
 PR 26-JUN-1998; 98US-0091010P.  
 PR 01-JUL-1998; 98US-0091359P.  
 PR 01-JUL-1998; 98US-0091544P.  
 PR 02-JUL-1998; 98US-0091478P.  
 PR 02-JUL-1998; 98US-0091486P.  
 PR 02-JUL-1998; 98US-0091626P.  
 PR 02-JUL-1998; 98US-0091628P.  
 PR 02-JUL-1998; 98US-0091632P.  
 PR 04-AUG-1998; 98US-0094006P.  
 PR 10-AUG-1998; 98US-0095998P.  
 PR 10-AUG-1998; 98US-0096012P.  
 PR 17-AUG-1998; 98US-0096757P.  
 PR 17-AUG-1998; 98US-0096766P.  
 PR 17-AUG-1998; 98US-0096867P.  
 PR 17-AUG-1998; 98US-0096891P.  
 PR 17-AUG-1998; 98US-0096897P.  
 PR 18-AUG-1998; 98US-0096949P.  
 PR 18-AUG-1998; 98US-0096959P.  
 PR 18-AUG-1998; 98US-0097022P.  
 PR 26-AUG-1998; 98US-0097952P.  
 PR 26-AUG-1998; 98US-0097954P.  
 PR 26-AUG-1998; 98US-0097955P.  
 PR 26-AUG-1998; 98US-0097971P.  
 PR 26-AUG-1998; 98US-0097974P.  
 PR 26-AUG-1998; 98US-0098014P.  
 PR 01-SEP-1998; 98US-0098716P.  
 PR 01-SEP-1998; 98US-0098723P.  
 PR 02-SEP-1998; 98US-0098803P.  
 PR 02-SEP-1998; 98US-0098821P.  
 PR 02-SEP-1998; 98US-0098843P.  
 PR 09-SEP-1998; 98US-0099602P.  
 PR 10-SEP-1998; 98US-0099741P.  
 PR 10-SEP-1998; 98US-0099754P.  
 PR 10-SEP-1998; 98US-0099763P.  
 PR 10-SEP-1998; 98US-0099812P.  
 PR 15-SEP-1998; 98US-0100388P.  
 PR 16-SEP-1998; 98US-0100662P.  
 PR 16-SEP-1998; 98US-0100664P.  
 PR 16-SEP-1998; 98US-0101751P.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 17-SEP-1998; 98US-0100683P.  
 PR 17-SEP-1998; 98US-0100684P.  
 PR 17-SEP-1998; 98US-0100919P.  
 PR 17-SEP-1998; 98US-0100930P.  
 PR 18-SEP-1998; 98US-0100849P.  
 PR 18-SEP-1998; 98US-0101014P.  
 PR 18-SEP-1998; 98US-0101068P.  
 PR 23-SEP-1998; 98US-0101471P.  
 PR 23-SEP-1998; 98US-0101472P.  
 PR 23-SEP-1998; 98US-0101475P.  
 PR 23-SEP-1998; 98US-0101477P.  
 PR 24-SEP-1998; 98US-0101738P.  
 PR 24-SEP-1998; 98US-0101739P.  
 PR 24-SEP-1998; 98US-0101743P.  
 PR 24-SEP-1998; 98US-0101922P.

```

PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.

Query Match 66.7%; Score 6; DB 6; Length 234;
Best Local Similarity 100.0%; Pred.No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWFL 6
Db 166 IETWFL 171

RESULT 16
ABU84295
ID ABU84295 standard; protein; 234 AA.
XX AC ABU84295;
XX XX
DT 02-AUG-2003 (first entry)
XX DE Human secreted/transmembrane protein (PRO) #33.
XX KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX OS Homo sapiens.
XX XX
XX PN US2003032112-A1.
XX PD 13-FEB-2003.
XX XX
XX PF 21-JUN-2002; 2002US-00176756.
XX XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-00633486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.

PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 28-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085800P.
PR 15-MAY-1998; 98US-0085862P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 03-JUN-1998; 98US-0087759P.
PR 04-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 17-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.

```

PR 25-JUN-1998;	98US-0090688P.	PR 01-OCT-1998;	98US-0102687P.
PR 25-JUN-1998;	98US-0090690P.	PR 02-OCT-1998;	98US-0102965P.
PR 25-JUN-1998;	98US-0090694P.	PR 06-OCT-1998;	98US-0103258P.
PR 25-JUN-1998;	98US-0090695P.	PR 06-OCT-1998;	98US-0103449P.
PR 26-JUN-1998;	98US-0090696P.	PR 07-OCT-1998;	98US-00168978.
PR 26-JUN-1998;	98US-0090862P.	Query Match 66.7%; Score 6; DB 6; Length 234;	
PR 26-JUN-1998;	98US-0090863P.	Best Local Similarity 100.0%; Pred. No. 26;	
PR 26-JUN-1998;	98US-0091010P.	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
PR 01-JUL-1998;	98US-0091359P.	OY 1 IETWEL 6	
PR 01-JUL-1998;	98US-0091544P.	Db 166 IETWEL 171	
PR 02-JUL-1998;	98US-0091478P.		
PR 02-JUL-1998;	98US-0091486P.	RESULT 17	
PR 02-JUL-1998;	98US-0091626P.	ABR66169	
PR 02-JUL-1998;	98US-0091628P.	ID ABR66169 standard; protein; 234 AA.	
PR 02-JUL-1998;	98US-0091632P.	XX AC ABR66169;	
PR 04-AUG-1998;	98US-0094006P.	XX DT 05-AUG-2003 (first entry)	
PR 10-AUG-1998;	98US-0095998P.	XX DE Human secreted polypeptide PRO1864, SEQ ID NO:66.	
PR 17-AUG-1998;	98US-0096757P.	XX KW Human; PRO; secreted protein; transmembrane protein;	
PR 17-AUG-1998;	98US-0096766P.	XX KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;	
PR 17-AUG-1998;	98US-0096897P.	XX KW Chondrocyte; proliferation; differentiation; cartilage disorder;	
PR 18-AUG-1998;	98US-0096949P.	XX KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;	
PR 18-AUG-1998;	98US-0096959P.	XX KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;	
PR 26-AUG-1998;	98US-0097952P.	XX KW liver; drug screening; transgenic animal; genetic analysis;	
PR 26-AUG-1998;	98US-0097955P.	XX KW antiarthritic; vulnery; gene therapy.	
PR 26-AUG-1998;	98US-0097971P.	XX OS Homo sapiens.	
PR 26-AUG-1998;	98US-0097974P.	XX PN US2003027278-A1.	
PR 26-AUG-1998;	98US-0098014P.	XX PD 06-FEB-2003.	
PR 01-SEP-1998;	98US-0098716P.	XX PF 21-JUN-2002; 2002US-00176987.	
PR 01-SEP-1998;	98US-0098723P.	XX PR 18-SEP-1997;	
PR 02-SEP-1998;	98US-0098803P.	XX PR 18-SEP-1997;	
PR 02-SEP-1998;	98US-0098821P.	XX PR 17-OCT-1997;	
PR 02-SEP-1998;	98US-0098843P.	XX PR 21-OCT-1997;	
PR 09-SEP-1998;	98US-0099602P.	XX PR 24-OCT-1997;	
PR 10-SEP-1998;	98US-0099741P.	XX PR 24-OCT-1997;	
PR 10-SEP-1998;	98US-0099754P.	XX PR 28-OCT-1997;	
PR 10-SEP-1998;	98US-0099763P.	XX PR 28-OCT-1997;	
PR 10-SEP-1998;	98US-0099812P.	XX PR 28-OCT-1997;	
PR 15-SEP-1998;	98US-0100388P.	XX PR 28-OCT-1997;	
PR 16-SEP-1998;	98US-0100662P.	XX PR 29-OCT-1997;	
PR 16-SEP-1998;	98US-0100664P.	XX PR 31-OCT-1997;	
PR 16-SEP-1998;	98US-0101751P.	XX PR 31-OCT-1997;	
PR 17-SEP-1998;	98US-0100683P.	XX PR 13-NOV-1997;	
PR 17-SEP-1998;	98US-0100684P.	XX PR 13-NOV-1997;	
PR 17-SEP-1998;	98US-0100930P.	XX PR 24-NOV-1997;	
PR 18-SEP-1998;	98US-0101014P.	XX PR 24-NOV-1997;	
PR 18-SEP-1998;	98US-0101068P.	XX PR 11-DEC-1997;	
PR 23-SEP-1998;	98US-0101471P.	XX PR 12-DEC-1997;	
PR 23-SEP-1998;	98US-0101472P.	XX PR 17-DEC-1997;	
PR 23-SEP-1998;	98US-0101475P.	XX PR 18-DEC-1997;	
PR 23-SEP-1998;	98US-0101777P.	XX PR 10-MAR-1998;	
PR 24-SEP-1998;	98US-0101738P.	XX PR 11-MAR-1998;	
PR 24-SEP-1998;	98US-0101739P.	XX PR 20-MAR-1998;	
PR 24-SEP-1998;	98US-0101743P.	XX PR 20-MAR-1998;	
PR 25-SEP-1998;	98US-0101822P.	XX PR 27-MAR-1998;	
PR 25-SEP-1998;	98US-0101786P.	XX PR 27-MAR-1998;	
PR 29-SEP-1998;	98US-0102207P.	XX PR 31-MAR-1998;	
PR 29-SEP-1998;	98US-0102240P.	XX PR 01-APR-1998;	
PR 29-SEP-1998;	98US-0102330P.	XX PR 01-APR-1998;	
PR 29-SEP-1998;	98US-0102331P.	XX PR 01-APR-1998;	
PR 30-SEP-1998;	98US-0102487P.	XX PR 01-APR-1998;	
PR 30-SEP-1998;	98US-0102570P.	XX PR 01-APR-1998;	
PR 30-SEP-1998;	98US-0102571P.	XX PR 01-APR-1998;	
PR 01-OCT-1998;	98US-0102684P.	XX PR 01-APR-1998;	

PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 28-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086332P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 09-JUN-1998; 98US-0088217P.  
PR 10-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.

PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095988P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 01-SEP-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 03-SEP-1998; 98US-0098843P.  
PR 03-SEP-1998; 98US-0098602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.

Query Match

66.7%; Score 6; DB 6; Length 234;

Best Local Similarity 100.0%; Pred. No. 26;		Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 IETWFL 6		
Db	166 IETWFL 171		
RESULT 18			
ABR65559			
ID	ABR65559 standard; protein; 234 AA.		
XX			
AC	ABR65559;		
XX			
DT	05-AUG-2003 (first entry)		
XX			
DE	Human secreted polypeptide PRO1864, SEQ ID NO:66.		
XX			
KW	Human; PRO; secreted protein; transmembrane protein; extracellular domain; tumour necrosis factor-alpha; TNF-alpha;		
KW	chondrocyte; proliferation; differentiation; cartilage disorder;		
KW	bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;		
KW	adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;		
KW	liver; drug screening; transgenic animal; genetic analysis;		
KW	antiarthritic; vulnery; gene therapy.		
XX			
OS	Homo sapiens.		
XX			
PN	US2003036159-A1.		
XX			
PD	20-FEB-2003.		
XX			
PF	02-JUL-2002; 2002US-00188773.		
XX			
PR	18-SEP-1997;	97US-0059263P.	98US-0082797P.
PR	18-SEP-1997;	97US-0059266P.	98US-0083322P.
PR	17-OCT-1997;	97US-0062250P.	98US-0083495P.
PR	21-OCT-1997;	97US-0063486P.	98US-0083498P.
PR	24-OCT-1997;	97US-0063120P.	98US-0083499P.
PR	24-OCT-1997;	97US-0063121P.	98US-0083559P.
PR	28-OCT-1997;	97US-0063540P.	98US-0084366P.
PR	28-OCT-1997;	97US-0063541P.	98US-0084414P.
PR	28-OCT-1997;	97US-0063544P.	98US-0084639P.
PR	28-OCT-1997;	97US-0063564P.	98US-0084640P.
PR	29-OCT-1997;	97US-0063734P.	98US-0084643P.
PR	31-OCT-1997;	97US-0063870P.	98US-0084644P.
PR	31-OCT-1997;	97US-0064103P.	98US-0084645P.
PR	13-NOV-1997;	97US-0065311P.	98US-0084646P.
PR	21-NOV-1997;	97US-0066120P.	98US-0084647P.
PR	24-NOV-1997;	97US-0066466P.	98US-0084648P.
PR	24-NOV-1997;	97US-0066772P.	98US-0084649P.
PR	11-DEC-1997;	97US-0069335P.	98US-0084650P.
PR	12-DEC-1997;	97US-0069425P.	98US-0084651P.
PR	17-DEC-1997;	97US-0069870P.	98US-0084652P.
PR	18-DEC-1997;	97US-0068017P.	98US-0084653P.
PR	10-MAR-1998;	98US-0077450P.	98US-0084654P.
PR	11-MAR-1998;	98US-0077632P.	98US-0084655P.
PR	11-MAR-1998;	98US-0077649P.	98US-0084656P.
PR	20-MAR-1998;	98US-0077886P.	98US-0084657P.
PR	20-MAR-1998;	98US-0078939P.	98US-0084658P.
PR	27-MAR-1998;	98US-0079664P.	98US-0084659P.
PR	31-MAR-1998;	98US-0079786P.	98US-0084660P.
PR	31-MAR-1998;	98US-0080107P.	98US-0084661P.
PR	31-MAR-1998;	98US-0080194P.	98US-0084662P.
PR	01-APR-1998;	98US-0080327P.	98US-0084663P.
PR	01-APR-1998;	98US-0080333P.	98US-0084664P.
PR	08-APR-1998;	98US-0081049P.	98US-0084665P.
PR	08-APR-1998;	98US-0081070P.	98US-0084666P.
PR	09-APR-1998;	98US-0081195P.	98US-0084667P.
PR	15-APR-1998;	98US-0081838P.	98US-0084668P.
PR	21-APR-1998;	98US-0082568P.	98US-0084669P.
PR	21-APR-1998;	98US-0082569P.	98US-0084670P.
PR	22-APR-1998;	98US-0082704P.	98US-0084671P.
PR	22-APR-1998;	98US-0082705P.	98US-0084672P.
PR	22-APR-1998;	98US-0082706P.	98US-0084673P.
PR	22-APR-1998;	98US-0082707P.	98US-0084674P.
PR	22-APR-1998;	98US-0082708P.	98US-0084675P.
PR	22-APR-1998;	98US-0082709P.	98US-0084676P.
PR	22-APR-1998;	98US-0082710P.	98US-0084677P.
PR	22-APR-1998;	98US-0082711P.	98US-0084678P.
PR	22-APR-1998;	98US-0082712P.	98US-0084679P.
PR	22-APR-1998;	98US-0082713P.	98US-0084680P.
PR	22-APR-1998;	98US-0082714P.	98US-0084681P.
PR	22-APR-1998;	98US-0082715P.	98US-0084682P.
PR	22-APR-1998;	98US-0082716P.	98US-0084683P.
PR	22-APR-1998;	98US-0082717P.	98US-0084684P.
PR	22-APR-1998;	98US-0082718P.	98US-0084685P.
PR	22-APR-1998;	98US-0082719P.	98US-0084686P.
PR	22-APR-1998;	98US-0082720P.	98US-0084687P.
PR	22-APR-1998;	98US-0082721P.	98US-0084688P.
PR	22-APR-1998;	98US-0082722P.	98US-0084689P.
PR	22-APR-1998;	98US-0082723P.	98US-0084690P.
PR	22-APR-1998;	98US-0082724P.	98US-0084691P.
PR	22-APR-1998;	98US-0082725P.	98US-0084692P.
PR	22-APR-1998;	98US-0082726P.	98US-0084693P.
PR	22-APR-1998;	98US-0082727P.	98US-0084694P.
PR	22-APR-1998;	98US-0082728P.	98US-0084695P.
PR	22-APR-1998;	98US-0082729P.	98US-0084696P.
PR	22-APR-1998;	98US-0082730P.	98US-0084697P.
PR	22-APR-1998;	98US-0082731P.	98US-0084698P.
PR	22-APR-1998;	98US-0082732P.	98US-0084699P.
PR	22-APR-1998;	98US-0082733P.	98US-0084700P.
PR	22-APR-1998;	98US-0082734P.	98US-0084701P.
PR	22-APR-1998;	98US-0082735P.	98US-0084702P.
PR	22-APR-1998;	98US-0082736P.	98US-0084703P.
PR	22-APR-1998;	98US-0082737P.	98US-0084704P.
PR	22-APR-1998;	98US-0082738P.	98US-0084705P.
PR	22-APR-1998;	98US-0082739P.	98US-0084706P.
PR	22-APR-1998;	98US-0082740P.	98US-0084707P.
PR	22-APR-1998;	98US-0082741P.	98US-0084708P.
PR	22-APR-1998;	98US-0082742P.	98US-0084709P.
PR	22-APR-1998;	98US-0082743P.	98US-0084710P.
PR	22-APR-1998;	98US-0082744P.	98US-0084711P.
PR	22-APR-1998;	98US-0082745P.	98US-0084712P.
PR	22-APR-1998;	98US-0082746P.	98US-0084713P.
PR	22-APR-1998;	98US-0082747P.	98US-0084714P.
PR	22-APR-1998;	98US-0082748P.	98US-0084715P.
PR	22-APR-1998;	98US-0082749P.	98US-0084716P.
PR	22-APR-1998;	98US-0082750P.	98US-0084717P.
PR	22-APR-1998;	98US-0082751P.	98US-0084718P.
PR	22-APR-1998;	98US-0082752P.	98US-0084719P.
PR	22-APR-1998;	98US-0082753P.	98US-0084720P.
PR	22-APR-1998;	98US-0082754P.	98US-0084721P.
PR	22-APR-1998;	98US-0082755P.	98US-0084722P.
PR	22-APR-1998;	98US-0082756P.	98US-0084723P.
PR	22-APR-1998;	98US-0082757P.	98US-0084724P.
PR	22-APR-1998;	98US-0082758P.	98US-0084725P.
PR	22-APR-1998;	98US-0082759P.	98US-0084726P.
PR	22-APR-1998;	98US-0082760P.	98US-0084727P.
PR	22-APR-1998;	98US-0082761P.	98US-0084728P.
PR	22-APR-1998;	98US-0082762P.	98US-0084729P.
PR	22-APR-1998;	98US-0082763P.	98US-0084730P.
PR	22-APR-1998;	98US-0082764P.	98US-0084731P.
PR	22-APR-1998;	98US-0082765P.	98US-0084732P.
PR	22-APR-1998;	98US-0082766P.	98US-0084733P.
PR	22-APR-1998;	98US-0082767P.	98US-0084734P.
PR	22-APR-1998;	98US-0082768P.	98US-0084735P.
PR	22-APR-1998;	98US-0082769P.	98US-0084736P.
PR	22-APR-1998;	98US-0082770P.	98US-0084737P.
PR	22-APR-1998;	98US-0082771P.	98US-0084738P.
PR	22-APR-1998;	98US-0082772P.	98US-0084739P.
PR	22-APR-1998;	98US-0082773P.	98US-0084740P.
PR	22-APR-1998;	98US-0082774P.	98US-0084741P.
PR	22-APR-1998;	98US-0082775P.	98US-0084742P.
PR	22-APR-1998;	98US-0082776P.	98US-0084743P.
PR	22-APR-1998;	98US-0082777P.	98US-0084744P.
PR	22-APR-1998;	98US-0082778P.	98US-0084745P.
PR	22-APR-1998;	98US-0082779P.	98US-0084746P.
PR	22-APR-1998;	98US-0082780P.	98US-0084747P.
PR	22-APR-1998;	98US-0082781P.	98US-0084748P.
PR	22-APR-1998;	98US-0082782P.	98US-0084749P.
PR	22-APR-1998;	98US-0082783P.	98US-0084750P.
PR	22-APR-1998;	98US-0082784P.	98US-0084751P.
PR	22-APR-1998;	98US-0082785P.	98US-0084752P.
PR	22-APR-1998;	98US-0082786P.	98US-0084753P.
PR	22-APR-1998;	98US-0082787P.	98US-0084754P.
PR	22-APR-1998;	98US-0082788P.	98US-0084755P.
PR	22-APR-1998;	98US-0082789P.	98US-0084756P.
PR	22-APR-1998;	98US-0082790P.	98US-0084757P.
PR	22-APR-1998;	98US-0082791P.	98US-0084758P.
PR	22-APR-1998;	98US-0082792P.	98US-0084759P.
PR	22-APR-1998;	98US-0082793P.	98US-0084760P.
PR	22-APR-1998;	98US-0082794P.	98US-0084761P.
PR	22-APR-1998;	98US-0082795P.	98US-0084762P.
PR	22-APR-1998;	98US-0082796P.	98US-0084763P.
PR	22-APR-1998;	98US-0082797P.	98US-0084764P.
PR	22-APR-1998;	98US-0082798P.	98US-0084765P.
PR	22-APR-1998;	98US-0082799P.	98US-0084766P.
PR	22-APR-1998;	98US-0082800P.	98US-0084767P.
PR	22-APR-1998;	98US-0082801P.	98US-0084768P.
PR	22-APR-1998;	98US-0082802P.	98US-0084769P.
PR	22-APR-1998;	98US-0082803P.	98US-0084770P.
PR	22-APR-1998;	98US-0082804P.	98US-0084771P.
PR	22-APR-1998;	98US-0082805P.	98US-0084772P.
PR	22-APR-1998;	98US-0082806P.	98US-0084773P.
PR	22-APR-1998;	98US-0082807P.	98US-0084774P.
PR	22-APR-1998;	98US-0082808P.	98US-0084775P.
PR	22-APR-1998;	98US-0082809P.	98US-0084776P.
PR	22-APR-1998;	98US-0082810P.	98US-0084777P.
PR	22-APR-1998;	98US-0082811P.	98US-0084778P.
PR	22-APR-1998;	98US-0082812P.	98US-0084779P.
PR	22-APR-1998;	98US-0082813P.	98US-0084780P.
PR	22-APR-1998;	98US-0082814P.	98US-0084781P.
PR	22-APR-1998;	98US-0082815P.	98US-0084782P.
PR	22-APR-1998;	98US-0082816P.	98US-0084783P.
PR	22-APR-1998;	98US-0082817P.	98US-0084784P.
PR	22-APR-1998;	98US-0082818P.	98US-0084785P.
PR	22-APR-1998;	98US-0082819P.	98US-0084786P.
PR	22-APR-1998;	98US-0082820P.	98US-0084787P.
PR	22-APR-1998;	98US-0082821P.	98US-0084788P.
PR	22-APR-1998;	98US-0082822P.	98US-0084789P.
PR	22-APR-1998;	98US-0082823P.	98US-0084790P.
PR	22-APR-1998;	98US-0082824P.	98US-0084791P.
PR	22-APR-1998;	98US-0082825P.	98US-0084792P.
PR	22-APR-1998;	98US-0082826P.	98US-0084793P.
PR	22-APR-1998;	98US-0082827P.	98US-0084794P.
PR	22-APR-1998;	98US-0082828P.	98US-0084795P.
PR	22-APR-1998;	98US-0082829P.	98US-0084796P.
PR	22-APR-1998;	98US-0082830P.	98US-0084797P.
PR	22-APR-1998;	98US-0082831P.	98US-0084798P.
PR	22-APR-1998;	98US-0082832P.	98US-0084799P.
PR	22-APR-1998;	98US-0082833P.	98US-0084800P.
PR	22-APR-1998;	98US-0082834P.	98US-0084801P.
PR	22-APR-1998;	98US-0082835P.	98US-0084802P.
PR	22-APR-1998;	98US-0082836P.	98US-0084803P.
PR	22-APR-1998;	98US-0082837P.	98US-0084804P.
PR	22-APR-1998;	98US-0082838P.	98US-0084805P.
PR	22-APR-1998;	98US-0082839P.	98US-0084806P.
PR	22-APR-1998;	98US-0082840P.	98US-0084807P.
PR	22-APR-1998;	98US-0082841P.	98US-0084808P.
PR	22-APR-1998;	98US-0082842P.	98US-0084809P.

PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 02-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096756P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019130.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100910P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 25-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 25-SEP-1998; 98US-0102207P.  
PR 25-SEP-1998; 98US-0102240P.  
PR 25-SEP-1998; 98US-0102330P.  
PR 25-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.

Query Match 66.7%; Score 6; DB 6; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6

Db 166 IETWFL 171

RESULT 19  
ABU99499  
ID ABU99499 standard; protein; 234 AA.  
XX  
AC ABU99499;  
XX  
DT 09-AUG-2003 (first entry)  
XX  
DE Human secreted/transmembrane protein (PRO) #33.  
XX  
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;  
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;  
XX tissue typing.  
XX  
OS Homo sapiens.  
XX  
PN US2003040070-A1.  
XX  
PD 27-FEB-2003.  
XX  
PF 27-JUN-2002; 2002US-00184627.  
XX  
PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083222P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083599P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.

PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 11-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 16-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 17-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 25-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090576P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090689P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.

PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-00168978.

Query Match 66.7%; Score 6; DB 6; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6  
Db 166 IETWFL 171

RESULT 20  
ABU82738  
ID ABU82738 standard; protein; 234 AA.  
XX  
AC ABU82738;  
XX





PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 01-SEP-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-0016897P.  
PR 07-OCT-1998; 98US-0103395P.

Query Match 66.7%; Score 6; DB 6; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6  
DB 166 IETWFL 171

RESULT 21  
ABU89859  
ID ABU89859 standard; protein; 234 AA.

XX AC ABU89859;  
XX DT 11-AUG-2003 (first entry)  
XX DE Novel human secreted and transmembrane protein PRO1864.  
XX KW Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;  
KW chondrocyte differentiation; tumour necrosis factor-alpha release;  
KW affinity purification.

XX OS Homo sapiens.  
XX PN US2003036147-A1.  
XX PD 20-FEB-2003.  
XX PF 02-JUL-2002; 2002US-00187741.  
XX PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 11-DEC-1997; 97US-0066772P.  
PR 12-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087755P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.

```

PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089588P.
PR 17-JUN-1998; 98US-0089593P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090441P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 25-JUN-1998; 98US-0090876P.
PR 25-JUN-1998; 98US-0090878P.
PR 25-JUN-1998; 98US-0090888P.
PR 25-JUN-1998; 98US-0090909P.
PR 25-JUN-1998; 98US-0090949P.
PR 25-JUN-1998; 98US-0090959P.
PR 26-JUN-1998; 98US-0090966P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 04-JUL-1998; 98US-0094006P.
PR 04-JUL-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 26-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 15-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 66.7%; Score 6; DB 6; Length 234;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6
Db 166 IETWFL 171

RESULT 22
ABR68108
ID ABR68108 standard; protein; 234 AA.
XX
AC ABR68108;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO1864, SEQ ID NO:66.
XX
KW Human; PRO; secreted protein; transmembrane protein; TNF-alpha;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX

```

PN	US2003027264-A1.		
XX			
PD	06-FEB-2003.		
XX			
PF	18-JUN-2002; 2002US-00174579.		
XX			
PR	18-SEP-1997; 97US-0059263P.	PR	05-JUN-1998; 98US-0088167P.
PR	18-SEP-1997; 97US-0059263P.	PR	05-JUN-1998; 98US-0088202P.
PR	17-OCT-1997; 97US-0062250P.	PR	05-JUN-1998; 98US-0088212P.
PR	21-OCT-1997; 97US-0063486P.	PR	05-JUN-1998; 98US-0088217P.
PR	24-OCT-1997; 97US-0063120P.	PR	09-JUN-1998; 98US-0088655P.
PR	24-OCT-1997; 97US-0063121P.	PR	10-JUN-1998; 98US-0088722P.
PR	28-OCT-1997; 97US-0063540P.	PR	10-JUN-1998; 98US-0088738P.
PR	28-OCT-1997; 97US-0063541P.	PR	10-JUN-1998; 98US-0088740P.
PR	28-OCT-1997; 97US-0063544P.	PR	10-JUN-1998; 98US-0088811P.
PR	28-OCT-1997; 97US-0063564P.	PR	10-JUN-1998; 98US-0088824P.
PR	29-OCT-1997; 97US-0063734P.	PR	10-JUN-1998; 98US-0088825P.
PR	31-OCT-1997; 97US-0063870P.	PR	10-JUN-1998; 98US-0088826P.
PR	31-OCT-1997; 97US-0064103P.	PR	11-JUN-1998; 98US-0088861P.
PR	13-NOV-1997; 97US-0065311P.	PR	11-JUN-1998; 98US-0088863P.
PR	21-NOV-1997; 97US-0066120P.	PR	11-JUN-1998; 98US-0088876P.
PR	24-NOV-1997; 97US-0066466P.	PR	12-JUN-1998; 98US-0089090P.
PR	24-NOV-1997; 97US-0066772P.	PR	12-JUN-1998; 98US-0089105P.
PR	11-DEC-1997; 97US-0069335P.	PR	16-JUN-1998; 98US-0089512P.
PR	12-DEC-1997; 97US-0069425P.	PR	16-JUN-1998; 98US-0089514P.
PR	17-DEC-1997; 97US-0069870P.	PR	17-JUN-1998; 98US-0089538P.
PR	18-DEC-1997; 97US-0068017P.	PR	17-JUN-1998; 98US-0089598P.
PR	10-MAR-1998; 98US-0077450P.	PR	17-JUN-1998; 98US-0089653P.
PR	11-MAR-1998; 98US-0077632P.	PR	18-JUN-1998; 98US-0089908P.
PR	11-MAR-1998; 98US-0077649P.	PR	19-JUN-1998; 98US-0089952P.
PR	20-MAR-1998; 98US-0078886P.	PR	22-JUN-1998; 98US-0090246P.
PR	20-MAR-1998; 98US-0078939P.	PR	22-JUN-1998; 98US-0090252P.
PR	27-MAR-1998; 98US-0079664P.	PR	22-JUN-1998; 98US-0090254P.
PR	27-MAR-1998; 98US-0079786P.	PR	24-JUN-1998; 98US-0090429P.
PR	31-MAR-1998; 98US-0080107P.	PR	24-JUN-1998; 98US-0090435P.
PR	31-MAR-1998; 98US-0080194P.	PR	24-JUN-1998; 98US-0090444P.
PR	01-APR-1998; 98US-0080327P.	PR	24-JUN-1998; 98US-0090461P.
PR	01-APR-1998; 98US-0080333P.	PR	24-JUN-1998; 98US-0090535P.
PR	08-APR-1998; 98US-0081049P.	PR	24-JUN-1998; 98US-0090540P.
PR	08-APR-1998; 98US-0081070P.	PR	25-JUN-1998; 98US-0090676P.
PR	09-APR-1998; 98US-0081195P.	PR	25-JUN-1998; 98US-0090678P.
PR	15-APR-1998; 98US-0081838P.	PR	25-JUN-1998; 98US-0090688P.
PR	21-APR-1998; 98US-0082568P.	PR	25-JUN-1998; 98US-0090690P.
PR	21-APR-1998; 98US-0082569P.	PR	25-JUN-1998; 98US-0090694P.
PR	22-APR-1998; 98US-0082704P.	PR	25-JUN-1998; 98US-0090695P.
PR	22-APR-1998; 98US-0082797P.	PR	25-JUN-1998; 98US-0090696P.
PR	28-APR-1998; 98US-0083322P.	PR	26-JUN-1998; 98US-00105413.
PR	29-APR-1998; 98US-0083495P.	PR	26-JUN-1998; 98US-0090862P.
PR	29-APR-1998; 98US-0083496P.	PR	26-JUN-1998; 98US-0090863P.
PR	29-APR-1998; 98US-0083499P.	PR	26-JUN-1998; 98US-0091010P.
PR	29-APR-1998; 98US-0083559P.	PR	01-JUL-1998; 98US-0091359P.
PR	05-MAY-1998; 98US-0084366P.	PR	01-JUL-1998; 98US-0091544P.
PR	06-MAY-1998; 98US-0084414P.	PR	02-JUL-1998; 98US-0091478P.
PR	07-MAY-1998; 98US-0084639P.	PR	02-JUL-1998; 98US-0091486P.
PR	07-MAY-1998; 98US-0084640P.	PR	02-JUL-1998; 98US-0091626P.
PR	07-MAY-1998; 98US-0084643P.	PR	02-JUL-1998; 98US-0091628P.
PR	15-MAY-1998; 98US-0085579P.	PR	02-JUL-1998; 98US-0091632P.
PR	15-MAY-1998; 98US-0085580P.	PR	24-JUL-1998; 98US-0094006P.
PR	15-MAY-1998; 98US-0085582P.	PR	04-AUG-1998; 98US-0095282P.
PR	15-MAY-1998; 98US-0085700P.	PR	10-AUG-1998; 98US-0095998P.
PR	18-MAY-1998; 98US-0086023P.	PR	10-AUG-1998; 98US-0096012P.
PR	22-MAY-1998; 98US-0086392P.	PR	17-AUG-1998; 98US-0096757P.
PR	22-MAY-1998; 98US-0086486P.	PR	17-AUG-1998; 98US-0096766P.
PR	28-MAY-1998; 98US-0087098P.	PR	17-AUG-1998; 98US-0096867P.
PR	28-MAY-1998; 98US-0087208P.	PR	17-AUG-1998; 98US-0096891P.
PR	02-JUN-1998; 98US-0087609P.	PR	17-AUG-1998; 98US-0096897P.
PR	02-JUN-1998; 98US-0087759P.	PR	18-AUG-1998; 98US-0096949P.
PR	03-JUN-1998; 98US-0087827P.	PR	18-AUG-1998; 98US-0096959P.
PR	04-JUN-1998; 98US-0088025P.	PR	26-AUG-1998; 98US-0097022P.
PR	04-JUN-1998; 98US-0088028P.	PR	26-AUG-1998; 98US-0097522P.
PR	04-JUN-1998; 98US-0088029P.	PR	26-AUG-1998; 98US-0097954P.
PR	04-JUN-1998; 98US-0088033P.	PR	26-AUG-1998; 98US-0097955P.
PR	04-JUN-1998; 98US-0088326P.	PR	26-AUG-1998; 98US-0097971P.
		PR	26-AUG-1998; 98US-0097974P.
		PR	26-AUG-1998; 98US-0098014P.
		PR	01-SEP-1998; 98US-0098716P.
		PR	01-SEP-1998; 98US-0098723P.
		PR	02-SEP-1998; 98US-0099803P.
		PR	02-SEP-1998; 98US-0099821P.

```

PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100583P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100949P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102584P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 66.7%; Score 6; DB 6; Length 234;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6
DB 166 IETWFL 171

RESULT 23
ABU96161
XX AC ABU96161;
XX DT 25-JUL-2003 (first entry)
XX DE
XX DE Novel human secreted and transmembrane protein PRO1864.
KW Human; secreted and transmembrane protein; PRO; transgenic animal;
KW knockout; chromosome identification; tissue typing; tumour;
KW chondrocyte proliferation; chondrocyte differentiation;
KW tumor necrosis factor-alpha release stimulator.
XX OS Homo sapiens.
XX PN US2003036144-A1.
XX PD 20-FEB-2003.
XX PF
XX PF 01-JUL-2002; 2002US-00187601.
XX PR
XX PR 18-SEP-1997; 97US-0059263P.
XX PR 18-SEP-1997; 97US-0059266P.
XX PR 17-OCT-1997; 97US-0062250P.
XX PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086192P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088126P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.

```

PR 10-JUN-1998;	98US-0088825P.	PR 16-SEP-1998;	98WO-US019330.
PR 10-JUN-1998;	98US-0088826P.	PR 17-SEP-1998;	98US-0100683P.
PR 11-JUN-1998;	98US-0088861P.	PR 17-SEP-1998;	98US-0100684P.
PR 11-JUN-1998;	98US-0088863P.	PR 17-SEP-1998;	98US-0100919P.
PR 11-JUN-1998;	98US-0088876P.	PR 17-SEP-1998;	98US-0100930P.
PR 12-JUN-1998;	98US-0083090P.	PR 18-SEP-1998;	98US-0100849P.
PR 12-JUN-1998;	98US-0089105P.	PR 18-SEP-1998;	98US-0101014P.
PR 16-JUN-1998;	98US-0089512P.	PR 18-SEP-1998;	98US-0101068P.
PR 16-JUN-1998;	98US-0089514P.	PR 23-SEP-1998;	98US-0101471P.
PR 17-JUN-1998;	98US-0089538P.	PR 23-SEP-1998;	98US-0101472P.
PR 17-JUN-1998;	98US-0089598P.	PR 23-SEP-1998;	98US-0101475P.
PR 17-JUN-1998;	98US-0089653P.	PR 23-SEP-1998;	98US-0101477P.
PR 18-JUN-1998;	98US-0089908P.	PR 24-SEP-1998;	98US-0101738P.
PR 19-JUN-1998;	98US-0089952P.	PR 24-SEP-1998;	98US-0101739P.
PR 22-JUN-1998;	98US-0090246P.	PR 24-SEP-1998;	98US-0101743P.
PR 22-JUN-1998;	98US-0090252P.	PR 24-SEP-1998;	98US-0101922P.
PR 22-JUN-1998;	98US-0090254P.	PR 25-SEP-1998;	98US-0101786P.
PR 24-JUN-1998;	98US-0090423P.	PR 29-SEP-1998;	98US-0102207P.
PR 24-JUN-1998;	98US-0090435P.	PR 29-SEP-1998;	98US-0102240P.
PR 24-JUN-1998;	98US-0090444P.	PR 29-SEP-1998;	98US-0102330P.
PR 24-JUN-1998;	98US-0090461P.	PR 29-SEP-1998;	98US-0102331P.
PR 24-JUN-1998;	98US-0090533P.	PR 30-SEP-1998;	98US-0102487P.
PR 24-JUN-1998;	98US-0090540P.	PR 30-SEP-1998;	98US-0102570P.
PR 25-JUN-1998;	98US-0090676P.	PR 30-SEP-1998;	98US-0102571P.
PR 25-JUN-1998;	98US-0090678P.	PR 01-OCT-1998;	98US-0102684P.
PR 25-JUN-1998;	98US-0090688P.	PR 01-OCT-1998;	98US-0102687P.
PR 25-JUN-1998;	98US-0090690P.	PR 02-OCT-1998;	98US-0102965P.
PR 25-JUN-1998;	98US-0090694P.	PR 06-OCT-1998;	98US-0103258P.
PR 25-JUN-1998;	98US-0090695P.	PR 06-OCT-1998;	98US-0103449P.
PR 25-JUN-1998;	98US-0090696P.		
PR 26-JUN-1998;	98US-00105413.		
PR 26-JUN-1998;	98US-0090862P.		
PR 26-JUN-1998;	98US-0090863P.		
PR 26-JUN-1998;	98US-0091010P.		
PR 01-JUL-1998;	98US-0091359P.		
PR 01-JUL-1998;	98US-0091544P.		
PR 02-JUL-1998;	98US-0091478P.		
PR 02-JUL-1998;	98US-0091486P.		
PR 02-JUL-1998;	98US-0091628P.		
PR 02-JUL-1998;	98US-0091628P.		
PR 02-JUL-1998;	98US-0091632P.		
PR 24-JUL-1998;	98US-0094006P.		
PR 04-AUG-1998;	98US-0095282P.		
PR 10-AUG-1998;	98US-0095998P.		
PR 10-AUG-1998;	98US-0096012P.		
PR 17-AUG-1998;	98US-0096757P.		
PR 17-AUG-1998;	98US-0096766P.		
PR 17-AUG-1998;	98US-0096867P.		
PR 17-AUG-1998;	98US-0096891P.		
PR 17-AUG-1998;	98US-0096897P.		
PR 18-AUG-1998;	98US-0096943P.		
PR 18-AUG-1998;	98US-0096959P.		
PR 18-AUG-1998;	98US-0097022P.		
PR 26-AUG-1998;	98US-0097952P.		
PR 26-AUG-1998;	98US-0097954P.		
PR 26-AUG-1998;	98US-0097955P.		
PR 26-AUG-1998;	98US-0097971P.		
PR 26-AUG-1998;	98US-0097974P.		
PR 26-AUG-1998;	98US-0098014P.		
PR 01-SEP-1998;	98US-0098716P.		
PR 01-SEP-1998;	98US-0098723P.		
PR 02-SEP-1998;	98US-0098803P.		
PR 02-SEP-1998;	98US-0098821P.		
PR 02-SEP-1998;	98US-0098843P.		
PR 09-SEP-1998;	98US-0099602P.		
PR 10-SEP-1998;	98US-0099741P.		
PR 10-SEP-1998;	98US-0099754P.		
PR 10-SEP-1998;	98US-0099763P.		
PR 10-SEP-1998;	98US-0099812P.		
PR 15-SEP-1998;	98US-0100388P.		
PR 16-SEP-1998;	98US-0100662P.		
PR 16-SEP-1998;	98US-0100664P.		
PR 16-SEP-1998;	98US-0101751P.		
<p>Query Match 66.7%; Score 6; DB 6; Length 234;</p> <p>Best Local Similarity 100.0%; Pred. No. 26;</p> <p>Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>			
Qy	1 IETWFL 6		
Db	166 IETWFL 171		
<p>RESULT 24</p> <p>ABU92592</p> <p>ID ABU92592 standard; protein; 234 AA.</p> <p>XX</p> <p>AC ABU92592;</p> <p>XX</p> <p>DT 18-JUL-2003 (first entry)</p> <p>XX</p> <p>DE Human secreted/transmembrane protein (PRO) #33.</p> <p>XX</p> <p>KW Human; secreted protein; transmembrane protein; PRO; tumour;</p> <p>KW proliferation; differentiation; chondrocyte cell; TNF-alpha;</p> <p>KW tumour necrosis factor-alpha; gene therapy.</p> <p>XX</p> <p>OS Homo sapiens.</p> <p>XX</p> <p>PN US2003036149-A1.</p> <p>XX</p> <p>PD 20-FEB-2003.</p> <p>XX</p> <p>PF 02-JUL-2002; 2002US-00187746.</p> <p>XX</p> <p>PR 18-SEP-1997; 97US-0059263P.</p> <p>PR 18-SEP-1997; 97US-0059266P.</p> <p>PR 17-OCT-1997; 97US-0062250P.</p> <p>PR 21-OCT-1997; 97US-0063486P.</p> <p>PR 24-OCT-1997; 97US-0063120P.</p> <p>PR 24-OCT-1997; 97US-0063121P.</p> <p>PR 28-OCT-1997; 97US-0063540P.</p> <p>PR 28-OCT-1997; 97US-0063541P.</p> <p>PR 28-OCT-1997; 97US-0063544P.</p> <p>PR 28-OCT-1997; 97US-0063564P.</p> <p>PR 29-OCT-1997; 97US-0063734P.</p> <p>PR 31-OCT-1997; 97US-0063870P.</p>			

PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-00661120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 28-APR-1998; 98US-0083495P.  
PR 28-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088555P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 24-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.





```
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090850P.
PR 25-JUN-1998; 98US-0090859P.
PR 25-JUN-1998; 98US-0090895P.
PR 25-JUN-1998; 98US-0090966P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 18-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100819P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.

PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 66.7%; Score 6; DB 6; Length 234;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6
Db 166 IETWFL 171

RESULT 26
AB002721
ID AB002721 standard; protein; 234 AA.
XX AC AB002721;
XX DT 09-AUG-2003 (first entry)
XX DE Human secreted/transmembrane protein (PRO) #33.
XX KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
XX KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
XX KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
XX KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX OS Homo sapiens.
XX PN US2003040062-A1.
XX PD 27-FEB-2003.
XX PF 25-JUN-2002; 2002US-00180545.
XX PR 18-SEP-1997; 97US-0059263P.
XX PR 18-SEP-1997; 97US-0059266P.
XX PR 17-OCT-1997; 97US-0062250P.
XX PR 21-OCT-1997; 97US-0063486P.
XX PR 24-OCT-1997; 97US-0063120P.
XX PR 24-OCT-1997; 97US-0063121P.
XX PR 28-OCT-1997; 97US-0063540P.
XX PR 28-OCT-1997; 97US-0063541P.
XX PR 28-OCT-1997; 97US-0063544P.
XX PR 28-OCT-1997; 97US-0063564P.
XX PR 29-OCT-1997; 97US-0063734P.
XX PR 31-OCT-1997; 97US-0063870P.
XX PR 31-OCT-1997; 97US-0064103P.
XX PR 13-NOV-1997; 97US-0065311P.
XX PR 21-NOV-1997; 97US-0066120P.
XX PR 24-NOV-1997; 97US-0066466P.
XX PR 24-NOV-1997; 97US-0066772P.
XX PR 11-DEC-1997; 97US-0069335P.
XX PR 12-DEC-1997; 97US-0069425P.
XX PR 17-DEC-1997; 97US-0069870P.
XX PR 18-DEC-1997; 97US-0068017P.
XX PR 10-MAR-1998; 98US-0077450P.
XX PR 11-MAR-1998; 98US-0077632P.
XX PR 11-MAR-1998; 98US-0077649P.
XX PR 20-MAR-1998; 98US-0078886P.
```

PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083222P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 18-MAY-1998; 98US-0085700P.  
PR 22-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 24-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 02-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 15-SEP-1998; 98US-0099812P.  
PR 16-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 18-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 23-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.

```

PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102985P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 66.7%; Score 6; DB 6; Length 234;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IETWFL 6
Db 166 IETWFL 171

RESULT 27
ABR74875
ID ABR74875 standard; protein; 234 AA.
XX
AC ABR74875;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human secreted polypeptide PRO1864, SEQ ID NO:66.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003040056-A1.
XX
PD 27-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176916.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 09-JUN-1998; 98US-0088217P.
PR 10-JUN-1998; 98US-0088555P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 11-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088961P.
PR 11-JUN-1998; 98US-0088963P.
PR 11-JUN-1998; 98US-0088967P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.

```



PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083455P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086033P.  
PR 22-MAY-1998; 98US-0086332P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 03-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089593P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090896P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.

PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 26-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 01-SEP-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 02-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 30-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.

Query Match 66.7%; Score 6; DB 6; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IETWFL 6

Db 166 IETWFL 171

RESULT 29

ABU85610

ID ABU85610 standard; protein; 234 AA.

XX AC ABU85610;

XX DT 02-JUL-2003 (first entry)

XX DE Human PRO polypeptide #33.

XX Human; PRO; secreted polypeptide; transmembrane polypeptide;

KW tumour necrosis factor alpha; TNF-alpha; chondrocyte cell; tumour;

KW cytostatic.

XX OS Homo sapiens.

XX US2003036140-A1.

XX PD 20-FEB-2003.

XX PF 01-JUL-2002; 2002US-00187588.

XX PR 26-JUN-1998; 98US-00105413.

PR 16-SEP-1998; 98WO-US019330.

PR 07-OCT-1998; 98US-00168978.

PR 07-OCT-1998; 98WO-US021141.

PR 06-NOV-1998; 98US-00187368.

PR 01-DEC-1998; 98WO-US025108.

PR 07-DEC-1998; 98US-00202054.

PR 03-MAR-1999; 99US-00254311.

PR 08-MAR-1999; 99WO-US005028.

PR 14-MAY-1999; 99US-00311832.

PR 14-MAY-1999; 99WO-US010733.

PR 02-JUN-1999; 99WO-US012252.

PR 25-AUG-1999; 99US-00380137.

PR 25-AUG-1999; 99US-00380138.

PR 25-AUG-1999; 99US-00380139.

PR 25-AUG-1999; 99US-00380142.

PR 01-SEP-1999; 99WO-US020111.

PR 15-SEP-1999; 99WO-US021090.

PR 18-OCT-1999; 99US-00403297.

PR 12-NOV-1999; 99US-00423844.

PR 01-DEC-1999; 99WO-US028301.

PR 30-DEC-1999; 99WO-US028551.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 18-FEB-2000; 2000WO-US004341.

PR 18-FEB-2000; 2000WO-US004342.

PR 22-FEB-2000; 2000WO-US004414.

PR 24-FEB-2000; 2000WO-US005004.

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005841.

PR 15-MAR-2000; 2000WO-US006884.

PR 30-MAR-2000; 2000WO-US008439.

PR 17-MAY-2000; 2000WO-US013705.

PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.

PR 28-JUL-2000; 2000WO-US020710.

PR 22-AUG-2000; 2000US-00644848.

PR 24-AUG-2000; 2000WO-US023328.

PR 18-SEP-2000; 2000US-00664610.

PR 18-SEP-2000; 2000US-00665350.

PR 08-NOV-2000; 2000US-00709238.

PR 08-NOV-2000; 2000WO-US030952.

PR 01-DEC-2000; 2000WO-US032678.

PR 20-DEC-2000; 2000US-00747259.

PR 20-DEC-2000; 2000WO-US034956.

PR 28-FEB-2001; 2001WO-US006520.

PR 22-MAR-2001; 2001US-00816744.

PR 10-MAY-2001; 2001US-00854208.

PR 10-MAY-2001; 2001US-00854280.

PR 25-MAY-2001; 2001US-00866028.

PR 01-JUN-2001; 2001WO-US017800.

PR 05-JUN-2001; 2001US-00874503.

PR 20-JUN-2001; 2001WO-US019692.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-JUL-2001; 2001WO-US021735.

PR 18-JUL-2001; 2001US-00908827.

PR 30-JUL-2001; 2001US-00918585.

PR 06-AUG-2001; 2001US-00924419.

PR 13-AUG-2001; 2001US-00929404.

PR 16-AUG-2001; 2001US-00931836.

PR 28-AUG-2001; 2001US-00941992.

PR 29-AUG-2001; 2001WO-US027099.

PR 04-SEP-2001; 2001US-00946374.

PR 15-JAN-2002; 2002US-00052586.

XX (GETH ) GENENTECH INC.

XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;

PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

XX WPI; 2003-332028/31.

DR N-PSDB; ACA72803.

XX Three hundred and five nucleic acids encoding PRO polypeptides, useful for the manufacture of a medicament for diagnosing or treating tumor.

XX Claim 11; Fig 66; 707pp; English.

XX The invention relates to human PRO polypeptides (secreted and transmembrane polypeptides) and the PRO polynucleotides encoding them. The invention also relates to a method for stimulating the release of tumour necrosis factor alpha (TNF-alpha) from human blood by contacting the blood with a sequence of the invention, a method for stimulating the proliferation or differentiation of chondrocyte cells by contacting the cells with a PRO polypeptide and a method for detecting the presence of a tumour in a mammal. The polypeptides and polynucleotides are useful for the manufacture of a medicament for diagnosing or treating a tumour in a mammal. Sequences ABU8578-ABU8582 represent human PRO polypeptides of the invention. Note: The sequence data for this patent is also available in electronic format from USPTO at seqdata.uspto.gov/sequence.html

XX Sequence 234 AA;

Query Match 66.7%; Score 6; DB 6; Length 234;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IETWFL 6

Db 166 IETWFL 171

RESULT 30

ABU98770

ID ABU98770 standard; protein; 234 AA.

XX AC ABU98770;

XX DT 01-AUG-2003 (first entry)

XX DE Novel human secreted and transmembrane protein PRO1864.

XX Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy;

KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;

KW colon tumour; breast tumour; prostate tumour; rectal tumour;

KW cervical tumour; liver tumour; TNF-alpha release;

KW tumour necrosis factor alpha release; chondrocyte cell proliferation;

KW chondrocyte cell differentiation; pharmaceutical; diagnostic; biosensor; bioreactor.

XX OS Homo sapiens.  
XX PN US2003013153-A1.  
XX PD 16-JAN-2003.  
XX PF 19-JUN-2002; 2002US-00175737.  
XX 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090689P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 04-AUG-1998; 98US-0094006P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098723P.  
 PR 02-SEP-1998; 98US-0098803P.  
 PR 02-SEP-1998; 98US-0098821P.  
 PR 02-SEP-1998; 98US-0098843P.  
 PR 09-SEP-1998; 98US-0099602P.  
 PR 10-SEP-1998; 98US-0099741P.  
 PR 10-SEP-1998; 98US-0099754P.  
 PR 10-SEP-1998; 98US-0099763P.  
 PR 10-SEP-1998; 98US-0099812P.  
 PR 15-SEP-1998; 98US-0100388P.  
 PR 16-SEP-1998; 98US-0100662P.  
 PR 16-SEP-1998; 98US-0100664P.  
 PR 16-SEP-1998; 98US-0101751P.  
 PR 16-SEP-1998; 98US-0101751P.  
 PR 17-SEP-1998; 98US-0100683P.  
 PR 17-SEP-1998; 98US-0100684P.  
 PR 17-SEP-1998; 98US-0100919P.  
 PR 17-SEP-1998; 98US-0100930P.  
 PR 18-SEP-1998; 98US-0100849P.  
 PR 18-SEP-1998; 98US-0101014P.  
 PR 18-SEP-1998; 98US-0101068P.  
 PR 23-SEP-1998; 98US-0101471P.  
 PR 23-SEP-1998; 98US-0101472P.  
 PR 23-SEP-1998; 98US-0101472P.  
 PR 23-SEP-1998; 98US-0101475P.  
 PR 23-SEP-1998; 98US-0101477P.  
 PR 24-SEP-1998; 98US-0101738P.  
 PR 24-SEP-1998; 98US-0101739P.  
 PR 24-SEP-1998; 98US-0101743P.  
 PR 24-SEP-1998; 98US-0101922P.  
 PR 25-SEP-1998; 98US-0101786P.  
 PR 29-SEP-1998; 98US-0102207P.  
 PR 29-SEP-1998; 98US-0102240P.  
 PR 29-SEP-1998; 98US-0102330P.  
 PR 29-SEP-1998; 98US-0102331P.  
 PR 30-SEP-1998; 98US-0102487P.  
 PR 30-SEP-1998; 98US-0102570P.  
 PR 30-SEP-1998; 98US-0102571P.  
 PR 01-OCT-1998; 98US-0102684P.  
 PR 01-OCT-1998; 98US-0102684P.  
 PR 01-OCT-1998; 98US-0102687P.

Query Match 66.7%; Score 6; DB 6; Length 234;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6

Db 166 IETWFL 171

RESULT 31  
 ABU97985  
 ID ABU97985 standard; protein; 234 AA.

AC ABU97985;

DT 30-JUL-2003 (first entry)

DE Novel human secreted and transmembrane protein PRO1864.

XX Human; secreted and transmembrane protein; PRO; cytotstatic; gene therapy;  
 KW Chondrocyte stimulator; tumour; adrenal tumour; lung tumour;  
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;  
 KW cervical tumour; liver tumour; chromosome identification.

XX Homo sapiens.

XX US2003017544-A1.

XX 23-JAN-2003.

XX 21-JUN-2002; 2002US-00176915.

XX 18-SEP-1997; 97US-0059263P.

PR 18-SEP-1997; 97US-0059266P.  
 PR 17-OCT-1997; 97US-0062250P.  
 PR 21-OCT-1997; 97US-0063486P.  
 PR 24-OCT-1997; 97US-0063120P.  
 PR 24-OCT-1997; 97US-0063121P.  
 PR 28-OCT-1997; 97US-0063540P.  
 PR 28-OCT-1997; 97US-0063541P.  
 PR 28-OCT-1997; 97US-0063544P.  
 PR 28-OCT-1997; 97US-0063564P.  
 PR 29-OCT-1997; 97US-0063734P.  
 PR 31-OCT-1997; 97US-0063870P.  
 PR 31-OCT-1997; 97US-0064103P.  
 PR 13-NOV-1997; 97US-0065311P.  
 PR 21-NOV-1997; 97US-0066120P.  
 PR 24-NOV-1997; 97US-0066468P.  
 PR 24-NOV-1997; 97US-0066772P.  
 PR 11-DEC-1997; 97US-0069335P.  
 PR 12-DEC-1997; 97US-0069425P.  
 PR 17-DEC-1997; 97US-0069870P.  
 PR 18-DEC-1997; 97US-0068017P.  
 PR 10-MAR-1998; 98US-0077450P.  
 PR 11-MAR-1998; 98US-0077632P.  
 PR 11-MAR-1998; 98US-0077649P.  
 PR 20-MAR-1998; 98US-0078886P.  
 PR 20-MAR-1998; 98US-0078939P.  
 PR 27-MAR-1998; 98US-0079664P.  
 PR 27-MAR-1998; 98US-0079786P.  
 PR 31-MAR-1998; 98US-0080107P.  
 PR 31-MAR-1998; 98US-0080194P.  
 PR 01-APR-1998; 98US-0080327P.  
 PR 01-APR-1998; 98US-0080333P.  
 PR 08-APR-1998; 98US-0081049P.  
 PR 08-APR-1998; 98US-0081070P.  
 PR 09-APR-1998; 98US-0081195P.  
 PR 15-APR-1998; 98US-0081838P.  
 PR 21-APR-1998; 98US-0082568P.  
 PR 21-APR-1998; 98US-0082569P.  
 PR 22-APR-1998; 98US-0082704P.  
 PR 22-APR-1998; 98US-0082797P.  
 PR 28-APR-1998; 98US-0083322P.  
 PR 29-APR-1998; 98US-0083495P.  
 PR 29-APR-1998; 98US-0083496P.  
 PR 29-APR-1998; 98US-0083499P.  
 PR 29-APR-1998; 98US-0083559P.  
 PR 05-MAY-1998; 98US-0084366P.  
 PR 06-MAY-1998; 98US-0084414P.  
 PR 07-MAY-1998; 98US-0084639P.  
 PR 07-MAY-1998; 98US-0084640P.  
 PR 07-MAY-1998; 98US-0084643P.  
 PR 15-MAY-1998; 98US-0085579P.  
 PR 15-MAY-1998; 98US-0085580P.  
 PR 15-MAY-1998; 98US-0085582P.  
 PR 15-MAY-1998; 98US-0085700P.  
 PR 18-MAY-1998; 98US-0086023P.  
 PR 22-MAY-1998; 98US-0086392P.  
 PR 22-MAY-1998; 98US-0086486P.  
 PR 28-MAY-1998; 98US-0087098P.  
 PR 28-MAY-1998; 98US-0087208P.  
 PR 02-JUN-1998; 98US-0087609P.  
 PR 02-JUN-1998; 98US-0087759P.  
 PR 03-JUN-1998; 98US-0087827P.  
 PR 04-JUN-1998; 98US-0088025P.  
 PR 04-JUN-1998; 98US-0088028P.  
 PR 04-JUN-1998; 98US-0088029P.  
 PR 04-JUN-1998; 98US-0088033P.  
 PR 04-JUN-1998; 98US-0088326P.  
 PR 05-JUN-1998; 98US-0088167P.  
 PR 05-JUN-1998; 98US-0088202P.  
 PR 05-JUN-1998; 98US-0088212P.  
 PR 05-JUN-1998; 98US-0088217P.  
 PR 09-JUN-1998; 98US-0088655P.  
 PR 10-JUN-1998; 98US-0088722P.  
 PR 10-JUN-1998; 98US-0088738P.



PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088865P.  
PR 12-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 04-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.

PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 17-SEP-1998; 98US-0100833P.  
PR 17-SEP-1998; 98US-0100833P.  
PR 17-SEP-1998; 98US-0100833P.  
PR 17-SEP-1998; 98US-0100833P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 25-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 02-OCT-1998; 98US-0102887P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.

Query Match 66.7%; Score 6; DB 6; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFL 6  
Db 166 IETWFL 171

## RESULT 32

ABU91691  
ID ABU91691 standard; protein; 234 AA.

XX  
AC ABU91691;

XX  
DT 11-AUG-2003 (first entry)

XX  
DE Novel human secreted and transmembrane protein PRO1864.

XX  
KW Human; gene therapy; chromosome identification; tissue typing.

XX  
OS Homo sapiens.

XX  
PN US2003027277-A1.

XX  
PD 06-FEB-2003.

XX  
PF 21-JUN-2002; 2002US-00176985.

XX  
PR 18-SEP-1997; 97US-0059263P.

PR 18-SEP-1997; 97US-0059266P.

PR 17-OCT-1997; 97US-0062250P.

PR 21-OCT-1997; 97US-0063486P.

PR 24-OCT-1997; 97US-0063120P.

PR 24-OCT-1997; 97US-0063121P.

PR 28-OCT-1997; 97US-0063540P.

PR 28-OCT-1997; 97US-0063541P.

PR 28-OCT-1997; 97US-0063544P.

PR 29-OCT-1997; 97US-0063734P.

PR	31-OCT-1997;	97US-0063870P.	PR	16-JUN-1998;	98US-0089512P.
PR	31-OCT-1997;	97US-0064103P.	PR	16-JUN-1998;	98US-0089514P.
PR	13-NOV-1997;	97US-0065311P.	PR	17-JUN-1998;	98US-0089538P.
PR	21-NOV-1997;	97US-0066120P.	PR	17-JUN-1998;	98US-0089598P.
PR	24-NOV-1997;	97US-0066466P.	PR	17-JUN-1998;	98US-0089653P.
PR	24-NOV-1997;	97US-0066772P.	PR	18-JUN-1998;	98US-0089908P.
PR	11-DEC-1997;	97US-0069335P.	PR	19-JUN-1998;	98US-0089952P.
PR	12-DEC-1997;	97US-0069425P.	PR	22-JUN-1998;	98US-0090246P.
PR	17-DEC-1997;	97US-0069870P.	PR	22-JUN-1998;	98US-0090252P.
PR	18-DEC-1997;	97US-0068017P.	PR	22-JUN-1998;	98US-0090254P.
PR	10-MAR-1998;	98US-0077450P.	PR	24-JUN-1998;	98US-0090429P.
PR	11-MAR-1998;	98US-0077632P.	PR	24-JUN-1998;	98US-0090435P.
PR	20-MAR-1998;	98US-0077649P.	PR	24-JUN-1998;	98US-0090444P.
PR	20-MAR-1998;	98US-0079886P.	PR	24-JUN-1998;	98US-0090461P.
PR	20-MAR-1998;	98US-0079939P.	PR	24-JUN-1998;	98US-0090535P.
PR	27-MAR-1998;	98US-0079664P.	PR	24-JUN-1998;	98US-0090540P.
PR	27-MAR-1998;	98US-0079786P.	PR	25-JUN-1998;	98US-0090676P.
PR	31-MAR-1998;	98US-0080107P.	PR	25-JUN-1998;	98US-0090678P.
PR	31-MAR-1998;	98US-0080194P.	PR	25-JUN-1998;	98US-0090688P.
PR	01-APR-1998;	98US-0080327P.	PR	25-JUN-1998;	98US-0090690P.
PR	01-APR-1998;	98US-0080333P.	PR	25-JUN-1998;	98US-0090694P.
PR	08-APR-1998;	98US-0081049P.	PR	25-JUN-1998;	98US-0090695P.
PR	08-APR-1998;	98US-0081070P.	PR	25-JUN-1998;	98US-0090696P.
PR	09-APR-1998;	98US-0081195P.	PR	26-JUN-1998;	98US-00105413.
PR	15-APR-1998;	98US-0081838P.	PR	26-JUN-1998;	98US-0090862P.
PR	21-APR-1998;	98US-0082568P.	PR	26-JUN-1998;	98US-0090863P.
PR	21-APR-1998;	98US-0082569P.	PR	26-JUN-1998;	98US-0091010P.
PR	22-APR-1998;	98US-0082704P.	PR	01-JUL-1998;	98US-0091353P.
PR	22-APR-1998;	98US-0082797P.	PR	01-JUL-1998;	98US-0091544P.
PR	28-APR-1998;	98US-0083322P.	PR	02-JUL-1998;	98US-0091478P.
PR	29-APR-1998;	98US-0083495P.	PR	02-JUL-1998;	98US-0091486P.
PR	29-APR-1998;	98US-0083496P.	PR	02-JUL-1998;	98US-0091626P.
PR	29-APR-1998;	98US-0083499P.	PR	02-JUL-1998;	98US-0091628P.
PR	29-APR-1998;	98US-0083559P.	PR	02-JUL-1998;	98US-0091632P.
PR	05-MAY-1998;	98US-0084366P.	PR	24-JUL-1998;	98US-0094006P.
PR	06-MAY-1998;	98US-0084414P.	PR	10-AUG-1998;	98US-0095282P.
PR	07-MAY-1998;	98US-0084639P.	PR	10-AUG-1998;	98US-0095998P.
PR	07-MAY-1998;	98US-0084640P.	PR	17-AUG-1998;	98US-0096012P.
PR	15-MAY-1998;	98US-0085579P.	PR	17-AUG-1998;	98US-0096766P.
PR	15-MAY-1998;	98US-0085580P.	PR	17-AUG-1998;	98US-0096867P.
PR	15-MAY-1998;	98US-0085582P.	PR	17-AUG-1998;	98US-0096891P.
PR	18-MAY-1998;	98US-0085700P.	PR	17-AUG-1998;	98US-0096897P.
PR	22-MAY-1998;	98US-0086023P.	PR	18-AUG-1998;	98US-0096949P.
PR	22-MAY-1998;	98US-0086392P.	PR	18-AUG-1998;	98US-0096959P.
PR	22-MAY-1998;	98US-0086486P.	PR	26-AUG-1998;	98US-0097022P.
PR	28-MAY-1998;	98US-0087098P.	PR	26-AUG-1998;	98US-0097952P.
PR	28-MAY-1998;	98US-0087208P.	PR	26-AUG-1998;	98US-0097955P.
PR	02-JUN-1998;	98US-0087759P.	PR	26-AUG-1998;	98US-0097971P.
PR	03-JUN-1998;	98US-0087827P.	PR	26-AUG-1998;	98US-0097974P.
PR	04-JUN-1998;	98US-0088025P.	PR	26-AUG-1998;	98US-0098014P.
PR	04-JUN-1998;	98US-0088028P.	PR	01-SEP-1998;	98US-0098716P.
PR	04-JUN-1998;	98US-0088029P.	PR	01-SEP-1998;	98US-0098723P.
PR	04-JUN-1998;	98US-0088033P.	PR	02-SEP-1998;	98US-0098803P.
PR	04-JUN-1998;	98US-0088326P.	PR	02-SEP-1998;	98US-0098821P.
PR	05-JUN-1998;	98US-0088167P.	PR	02-SEP-1998;	98US-0098843P.
PR	05-JUN-1998;	98US-0088202P.	PR	09-SEP-1998;	98US-0098602P.
PR	05-JUN-1998;	98US-0088212P.	PR	10-SEP-1998;	98US-0099741P.
PR	05-JUN-1998;	98US-0088217P.	PR	10-SEP-1998;	98US-0099754P.
PR	09-JUN-1998;	98US-0088655P.	PR	10-SEP-1998;	98US-0099763P.
PR	10-JUN-1998;	98US-0088722P.	PR	10-SEP-1998;	98US-0099812P.
PR	10-JUN-1998;	98US-0088738P.	PR	15-SEP-1998;	98US-0100388P.
PR	10-JUN-1998;	98US-0088740P.	PR	16-SEP-1998;	98US-0100662P.
PR	10-JUN-1998;	98US-0088811P.	PR	16-SEP-1998;	98US-0100664P.
PR	10-JUN-1998;	98US-0088824P.	PR	16-SEP-1998;	98US-0101751P.
PR	10-JUN-1998;	98US-0088825P.	PR	17-SEP-1998;	98WO-US019330.
PR	10-JUN-1998;	98US-0088826P.	PR	17-SEP-1998;	98US-0100683P.
PR	11-JUN-1998;	98US-0088861P.	PR	17-SEP-1998;	98US-0100684P.
PR	11-JUN-1998;	98US-0088863P.	PR	17-SEP-1998;	98US-0100919P.
PR	11-JUN-1998;	98US-0088876P.	PR	17-SEP-1998;	98US-0100930P.
PR	12-JUN-1998;	98US-0089090P.	PR	18-SEP-1998;	98US-0100849P.
PR	12-JUN-1998;	98US-0089105P.	PR	18-SEP-1998;	98US-0101014P.

```

PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-01038978.
PR 07-OCT-1998; 98US-0103335P.
PR 07-OCT-1998; 98US-0103401P.

Query Match 66.7%; Score 6; DB 6; Length 234;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWFL 6
Db 166 IETWFL 171

RESULT 33
ABU89384
ID ABU89384 standard; protein; 234 AA.
XX
AC ABU89384;
XX
DT 09-JUL-2003 (first entry)
XX
DE Human PRO polypeptide #33.
XX
KW Human; PRO polypeptide; secreted protein; transmembrane protein;
KW chromosome mapping; gene mapping; tumour; adrenal; lung; colon; breast;
KW prostate; rectal; cervical; liver; cancer; TNF-alpha;
KW tumour necrosis factor-alpha; proliferation; differentiation;
KW chondrocyte cell; bone disorder; cartilage disorder; sports injury;
KW arthritis; cytostatic; antiarthritic; osteopathic.
XX
OS Homo sapiens.
XX
XX US2003036141-A1.
XX
XX 20-FEB-2003.
XX
XX 01-JUL-2002; 2002US-00187597.
XX
XX 18-SEP-1997; 97US-0059263P.
XX 18-SEP-1997; 97US-0059266P.
XX 17-OCT-1997; 97US-0062250P.
XX 21-OCT-1997; 97US-0063486P.
XX 24-OCT-1997; 97US-0063120P.
XX 24-OCT-1997; 97US-0063121P.
XX 28-OCT-1997; 97US-0063540P.
XX 28-OCT-1997; 97US-0063541P.
XX 28-OCT-1997; 97US-0063544P.
XX 28-OCT-1997; 97US-0063564P.
XX 23-OCT-1997; 97US-0063734P.
XX 31-OCT-1997; 97US-0063870P.
XX 31-OCT-1997; 97US-0064103P.

97US-0065311P.
97US-0066120P.
97US-0066466P.
97US-0066772P.
97US-0069335P.
97US-0069425P.
97US-0069870P.
97US-0068017P.
98US-0077450P.
98US-0077632P.
98US-0077649P.
98US-0078866P.
98US-0078939P.
98US-0079664P.
98US-0079786P.
98US-0080107P.
98US-0080194P.
98US-0080327P.
98US-0080333P.
98US-0081049P.
98US-0081070P.
98US-0081195P.
98US-0081838P.
98US-0082568P.
98US-0082569P.
98US-0082704P.
98US-0082797P.
98US-0083222P.
98US-0083495P.
98US-0083496P.
98US-0083499P.
98US-0083559P.
98US-0084366P.
98US-0084414P.
98US-0084639P.
98US-0084640P.
98US-0084643P.
98US-0085579P.
98US-0085580P.
98US-0085582P.
98US-0085700P.
98US-0086023P.
98US-0086392P.
98US-0086486P.
98US-0087098P.
98US-0087208P.
98US-0087609P.
98US-0087759P.
98US-0087827P.
98US-0088025P.
98US-0088028P.
98US-0088029P.
98US-0088033P.
98US-0088326P.
98US-0088167P.
98US-0088202P.
98US-0088212P.
98US-0088217P.
98US-0088655P.
98US-0088722P.
98US-0088738P.
98US-0088740P.
98US-0088811P.
98US-0088824P.
98US-0088825P.
98US-0088826P.
98US-0088861P.
98US-0088863P.
98US-008876P.
98US-0089090P.
98US-0089105P.
98US-0089512P.
98US-0089514P.

```

PR 17-JUN-1998;	98US-0089538P.	PR 23-SEP-1998;	98US-0101472P.
PR 17-JUN-1998;	98US-0089598P.	PR 23-SEP-1998;	98US-0101475P.
PR 17-JUN-1998;	98US-0089653P.	PR 23-SEP-1998;	98US-0101477P.
PR 18-JUN-1998;	98US-0089908P.	PR 24-SEP-1998;	98US-0101738P.
PR 19-JUN-1998;	98US-0089952P.	PR 24-SEP-1998;	98US-0101739P.
PR 22-JUN-1998;	98US-0090246P.	PR 24-SEP-1998;	98US-0101743P.
PR 22-JUN-1998;	98US-0090252P.	PR 24-SEP-1998;	98US-0101922P.
PR 22-JUN-1998;	98US-0090254P.	PR 25-SEP-1998;	98US-0101786P.
PR 24-JUN-1998;	98US-0090429P.	PR 29-SEP-1998;	98US-0102207P.
PR 24-JUN-1998;	98US-0090435P.	PR 29-SEP-1998;	98US-0102240P.
PR 24-JUN-1998;	98US-0090444P.	PR 29-SEP-1998;	98US-0102330P.
PR 24-JUN-1998;	98US-0090461P.	PR 30-SEP-1998;	98US-0102331P.
PR 24-JUN-1998;	98US-0090535P.	PR 30-SEP-1998;	98US-0102487P.
PR 24-JUN-1998;	98US-0090540P.	PR 30-SEP-1998;	98US-0102570P.
PR 25-JUN-1998;	98US-0090676P.	PR 30-SEP-1998;	98US-0102571P.
PR 25-JUN-1998;	98US-0090678P.	PR 01-OCT-1998;	98US-0102684P.
PR 25-JUN-1998;	98US-0090688P.	PR 01-OCT-1998;	98US-0102687P.
PR 25-JUN-1998;	98US-0090690P.	PR 02-OCT-1998;	98US-0102965P.
PR 25-JUN-1998;	98US-0090694P.	Query Match 66.7%; Score 6; DB 6; Length 234;	
PR 25-JUN-1998;	98US-0090695P.	Best Local Similarity 100.0%; Pred.No. 26;	
PR 25-JUN-1998;	98US-0090696P.	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
PR 26-JUN-1998;	98US-00105413.	OY 1 IETWFL 6	
PR 26-JUN-1998;	98US-0090862P.	Db 166 IETWFL 171	
PR 26-JUN-1998;	98US-0090863P.		
PR 26-JUN-1998;	98US-0091010P.	RESULT 34	
PR 01-JUL-1998;	98US-0091359P.	ABU6225	
PR 01-JUL-1998;	98US-0091344P.	ID ABU6225 standard; protein; 234 AA.	
PR 02-JUL-1998;	98US-0091478P.	XX AC ABU6225;	
PR 02-JUL-1998;	98US-0091486P.	XX DT 01-JUL-2003 (first entry)	
PR 02-JUL-1998;	98US-0091626P.	XX DE Human secreted/transmembrane protein (PRO) #33.	
PR 02-JUL-1998;	98US-0091628P.	XX DE Human; immunogen; secreted protein; transmembrane protein; PRO; tumour;	
PR 02-JUL-1998;	98US-0091632P.	XX KW proliferation; differentiation; chondrocyte cells;	
PR 24-JUL-1998;	98US-0094006P.	XX KW tumour necrosis factor-alpha; TNF-alpha; blood; gene therapy.	
PR 04-AUG-1998;	98US-0095282P.	XX OS Homo sapiens.	
PR 10-AUG-1998;	98US-0095398P.	XX PN US2003036146-A1.	
PR 10-AUG-1998;	98US-0096012P.	XX PD 20-FEB-2003.	
PR 17-AUG-1998;	98US-0096757P.	XX PF 02-JUL-2002; 2002US-00187603.	
PR 17-AUG-1998;	98US-0096766P.	XX PR 26-JUN-1998; 98US-00105413.	
PR 17-AUG-1998;	98US-0096891P.	PR 16-SEP-1998; 98WO-US019330.	
PR 17-AUG-1998;	98US-0096897P.	PR 07-OCT-1998; 98US-00168978.	
PR 18-AUG-1998;	98US-0096949P.	PR 07-OCT-1998; 98WO-US021141.	
PR 18-AUG-1998;	98US-0096959P.	PR 06-NOV-1998; 98US-00187368.	
PR 18-AUG-1998;	98US-0097022P.	PR 01-DEC-1998; 98WO-US025108.	
PR 26-AUG-1998;	98US-0097952P.	PR 07-DEC-1998; 98US-00202054.	
PR 26-AUG-1998;	98US-0097954P.	PR 03-MAR-1999; 99US-00254311.	
PR 26-AUG-1998;	98US-0097971P.	PR 08-MAR-1999; 99WO-US005028.	
PR 26-AUG-1998;	98US-0097974P.	PR 14-MAY-1999; 99US-00311832.	
PR 26-AUG-1998;	98US-0098014P.	PR 14-MAY-1999; 99WO-US010733.	
PR 01-SEP-1998;	98US-009816P.	PR 02-JUN-1999; 99WO-US012252.	
PR 01-SEP-1998;	98US-0098723P.	PR 25-AUG-1999; 98US-00380137.	
PR 02-SEP-1998;	98US-0098803P.	PR 25-AUG-1999; 99US-00380139.	
PR 02-SEP-1998;	98US-0098843P.	PR 25-AUG-1999; 99US-00380142.	
PR 09-SEP-1998;	98US-0099602P.	PR 15-SEP-1999; 99WO-US020111.	
PR 10-SEP-1998;	98US-0099741P.	PR 18-OCT-1999; 99US-00403297.	
PR 10-SEP-1998;	98US-0099763P.	PR 12-NOV-1999; 99US-00423844.	
PR 10-SEP-1998;	98US-0099812P.	PR 01-DEC-1999; 99WO-US028301.	
PR 15-SEP-1998;	98US-0100388P.	PR 02-DEC-1999; 99WO-US028551.	
PR 16-SEP-1998;	98US-0100662P.	PR 30-DEC-1999; 99WO-US031274.	
PR 16-SEP-1998;	98US-0100664P.		
PR 16-SEP-1998;	98US-0101751P.		
PR 16-SEP-1998;	98WO-US019330.		
PR 17-SEP-1998;	98US-0100683P.		
PR 17-SEP-1998;	98US-0100684P.		
PR 17-SEP-1998;	98US-0100919P.		
PR 18-SEP-1998;	98US-0100849P.		
PR 18-SEP-1998;	98US-0101014P.		
PR 18-SEP-1998;	98US-0101068P.		
PR 23-SEP-1998;	98US-0101471P.		

05-JAN-2000; 2000WO-US0000219.  
 18-FEB-2000; 2000WO-US004341.  
 18-FEB-2000; 2000WO-US004342.  
 22-FEB-2000; 2000WO-US004414.  
 24-FEB-2000; 2000WO-US005004.  
 01-MAR-2000; 2000WO-US005001.  
 02-MAR-2000; 2000WO-US005841.  
 15-MAR-2000; 2000WO-US006884.  
 30-MAR-2000; 2000WO-US013705.  
 17-MAY-2000; 2000WO-US013705.  
 22-MAY-2000; 2000WO-US014042.  
 30-MAY-2000; 2000WO-US014941.  
 02-JUN-2000; 2000WO-US015264.  
 28-JUL-2000; 2000WO-US020710.  
 22-AUG-2000; 2000US-00644848.  
 24-AUG-2000; 2000WO-US023328.  
 18-SEP-2000; 2000US-00664610.  
 18-SEP-2000; 2000US-00665350.  
 08-NOV-2000; 2000US-00709238.  
 08-NOV-2000; 2000WO-US030952.  
 01-DEC-2000; 2000WO-US032678.  
 20-DEC-2000; 2000US-00747259.  
 20-DEC-2000; 2000WO-US034956.  
 28-FEB-2001; 2001WO-US006520.  
 22-MAR-2001; 2001US-00816744.  
 10-MAY-2001; 2001US-00854208.  
 10-MAY-2001; 2001US-00854280.  
 25-MAY-2001; 2001US-00866028.  
 01-JUN-2001; 2001WO-US017800.  
 05-JUN-2001; 2001US-00874503.  
 20-JUN-2001; 2001WO-US019692.  
 29-JUN-2001; 2001WO-US021066.  
 09-JUL-2001; 2001WO-US021735.  
 18-JUL-2001; 2001US-00908827.  
 30-JUL-2001; 2001US-00918585.  
 06-AUG-2001; 2001US-00924419.  
 13-AUG-2001; 2001US-00929404.  
 16-AUG-2001; 2001US-00931836.  
 28-AUG-2001; 2001US-00941992.  
 29-AUG-2001; 2001WO-US027099.  
 04-SEP-2001; 2001US-00946374.  
 15-JAN-2002; 2002US-00052586.  
 (GETH ) GENENTECH INC.  
 Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
 Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
 WPI; 2003-332034/31.  
 N-PSDB; ACA73417.  
 Three hundred and five nucleic acids encoding PRO polypeptides, useful in gene therapy, chromosome identification, tissue typing, and for detecting the presence of tumor in a mammal.  
 Claim 11; Fig 66; 707pp; English.  
 The invention relates to three hundred and five nucleic acids encoding PRO polypeptides (secreted and transmembrane), sequences 80% identical to them, or encoding a PRO polypeptide lacking its associated signal peptide or an extracellular domain of the PRO polypeptide, with or lacking its associated signal peptide. Also included are the encoded PRO proteins, PRO expression vectors, host cells transformed with the vector (used to produce PRO proteins), a chimaeric molecule comprising the PRO polypeptide fused to a heterologous amino acid sequence, an anti-PRO antibody, a method for stimulating the release of tumor necrosis factor alpha (TNF-alpha) from human blood (by contacting the blood with PRO1079, PRO827, PRO791, PRO1316, PRO1183, PRO1343, PRO1760, PRO1567 or PRO4333), a method for stimulating the proliferation or differentiation of chondrocyte cells by contacting the cells with a PRO6029 polypeptide, a method for detecting the presence of tumor in a mammal and an oligonucleotide probe derived from any of the nucleotide sequences cited above. The PRO polypeptide or anti-PRO antibody is useful for preparing a

CC medicament for treating a condition that is responsive to the PRO  
 CC polypeptide or anti-PRO antibody. The PRO nucleotide sequences are useful  
 CC as hybridisation probes in chromosome and gene mapping, or in generating  
 CC antisense RNA and DNA. PRO nucleic acids are also useful in preparing PRO  
 CC polypeptides, in assays to identify other proteins or molecules involved  
 CC in a binding reaction, to generate transgenic animals or knockout  
 CC animals, which in turn are useful in the development and screening of  
 CC therapeutically useful reagents, for chromosome identification, and  
 CC tissue typing. The PRO polypeptides and nucleic acid molecules are also  
 CC useful for detecting the presence of a tumour in a mammal, stimulating the  
 CC proliferation or differentiation of chondrocyte cells, stimulating the  
 CC release of tumour necrosis factor-alpha from human blood, in gene  
 CC therapy, or as molecular weight markers for protein electrophoresis  
 CC purposes. The anti-PRO antibodies may be used in diagnostic assays for  
 CC PRO, or for the affinity purification of PRO from recombinant cell  
 CC culture or natural sources. The present sequence represents a PRO protein  
 XX  
 SQ Sequence 234 AA;  
 Query Match 66.7%; Score 6; DB 6; Length 234;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IETWFL 6  
 Db 166 IETWFL 171  
 RESULT 35  
 ABU67438  
 ID ABU67438 standard; protein; 234 AA.  
 XX  
 AC ABU67438;  
 XX  
 DT 29-MAY-2003 (first entry)  
 XX  
 DE Human secreted/transmembrane protein (PRO) #33.  
 KW Human; secreted and transmembrane protein; PRO; TNF-alpha;  
 KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;  
 KW tissue typing.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003036162-A1.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 12-JUL-2002; 2002US-00194423.  
 XX  
 PR 26-JUN-1998; 98US-00105413.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 07-OCT-1998; 98US-00168978.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 06-NOV-1998; 98US-00187368.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 07-DEC-1998; 98US-00202054.  
 PR 03-MAR-1999; 98US-00254311.  
 PR 08-MAR-1999; 98WO-US005028.  
 PR 14-MAY-1999; 99US-00311832.  
 PR 14-MAY-1999; 99WO-US010733.  
 PR 02-JUN-1999; 99WO-US012252.  
 PR 25-AUG-1999; 99US-00380137.  
 PR 25-AUG-1999; 99US-00380138.  
 PR 25-AUG-1999; 99US-00380139.  
 PR 25-AUG-1999; 99US-00380142.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 18-OCT-1999; 99US-00403297.  
 PR 12-NOV-1999; 99US-00423844.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.  
PR 18-FEB-2000; 2000WO-US0004341.  
PR 18-FEB-2000; 2000WO-US0004342.  
PR 22-FEB-2000; 2000WO-US0004414.  
PR 24-FEB-2000; 2000WO-US0005004.  
PR 01-MAR-2000; 2000WO-US0005841.  
PR 02-MAR-2000; 2000WO-US0006884.  
PR 15-MAR-2000; 2000WO-US0008439.  
PR 30-MAR-2000; 2000WO-US013705.  
PR 17-MAY-2000; 2000WO-US014042.  
PR 22-MAY-2000; 2000WO-US014941.  
PR 30-MAY-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 22-AUG-2000; 2000US-00644848.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 18-SEP-2000; 2000US-00664610.  
PR 18-SEP-2000; 2000US-00665350.  
PR 08-NOV-2000; 2000US-00709238.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 22-MAR-2001; 2001US-00816744.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001US-00866028.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 30-JUL-2001; 2001US-00918585.  
PR 06-AUG-2001; 2001US-00924419.  
PR 13-AUG-2001; 2001US-00929404.  
PR 16-AUG-2001; 2001US-00931836.  
PR 28-AUG-2001; 2001US-00941992.  
PR 29-AUG-2001; 2001WO-US027099.  
PR 04-SEP-2001; 2001US-00946374.  
PR 15-JAN-2002; 2002US-00052586.

(GETH ) GENENTECH INC.

PA Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

XX WPI; 2003-332039/31.

XX N-PSDB; ACA05732.

XX New secreted and transmembrane PRO polypeptides and nucleic acids, useful  
PT in gene therapy, in chromosome and gene mapping, as chromosome markers,  
PT in tissue typing, and in chromosome identification.

XX Claim 11; Fig 66; 706pp; English.

XX The invention discloses human nucleic acids encoding secreted and  
CC transmembrane (PRO) polypeptides. Also disclosed is an antibody that  
CC specifically binds to the PRO polypeptide, a method for stimulating the  
CC release of tumour necrosis factor alpha (TNF-alpha) from human blood by  
CC contacting the blood a PRO polypeptide, a method for stimulating the  
CC proliferation or differentiation of chondrocyte cells by contacting the  
CC cells with a PRO polypeptide, a method for detecting the presence of a  
CC tumour in a mammal and an oligonucleotide probe derived from any of the  
CC PRO nucleotide sequences. The nucleotide sequences are useful as probes,  
CC in chromosome and gene mapping, in generating antisense RNA and DNA, in  
CC preparing PRO polypeptides by recombinant techniques and in gene therapy  
CC (e.g. for replacement of defective gene). The PRO polypeptides are useful  
CC as molecular weight markers for protein electrophoresis purposes, for  
CC chromosome identification, as chromosome markers, as therapeutic agents,  
CC for stimulating the release of TNF-alpha from human blood, for  
CC stimulating the proliferation or differentiation of chondrocytes and

CC detecting the presence of a tumour. The PRO polypeptides and nucleic  
CC acids may also be used diagnostically for tissue typing. The sequences  
CC presented in ABU67406-ABU67710 are the PRO polypeptides of the invention  
XX  
SQ Sequence 234 AA;

Query Match 66.7%; Score 6; DB 6; Length 234;  
Best Local Similarity 100.0%; Pred.No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IETWFL 6  
Db 166 IETWFL 171

RESULT 36

ABU80466  
ID ABU80466 standard; protein; 234 AA.

XX AC ABU80466;

XX DT 23-JUN-2003 (first entry)

XX DE Human PRO protein #33.

KW Human; tumour; adrenal; lung; colon; breast; prostate; rectal; cervical;  
KW liver; PRO; gene therapy.

XX OS Homo sapiens.

XX PN US2003036137-A1.

PD 20-FEB-2003.

XX PF 27-JUN-2002; 2002US-00184640.

XX PR 26-JUN-1998; 98US-00105413.

PR 16-SEP-1998; 98WO-US019330.

PR 07-OCT-1998; 98US-00168978.

PR 07-OCT-1998; 98WO-US021141.

PR 06-NOV-1998; 98US-00187368.

PR 01-DEC-1998; 98WO-US025108.

PR 07-DEC-1998; 98US-00202054.

PR 03-MAR-1999; 99US-00254311.

PR 08-MAR-1999; 99WO-US005028.

PR 14-MAY-1999; 99US-00311832.

PR 14-MAY-1999; 99WO-US010733.

PR 02-JUN-1999; 99WO-US012252.

PR 25-AUG-1999; 99US-00380137.

PR 25-AUG-1999; 99US-00380139.

PR 25-AUG-1999; 99US-00380142.

PR 01-SEP-1999; 99WO-US020111.

PR 15-SEP-1999; 99WO-US021090.

PR 18-OCT-1999; 99US-00403297.

PR 12-NOV-1999; 99US-00423844.

PR 01-DEC-1999; 99WO-US028301.

PR 02-DEC-1999; 99WO-US028551.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 18-FEB-2000; 2000WO-US004341.

PR 18-FEB-2000; 2000WO-US004342.

PR 22-FEB-2000; 2000WO-US004414.

PR 24-FEB-2000; 2000WO-US005004.

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005841.

PR 15-MAR-2000; 2000WO-US006884.

PR 30-MAR-2000; 2000WO-US008439.

PR 17-MAY-2000; 2000WO-US013705.

PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.

PR 28-JUL-2000; 2000WO-US020710.



XX 10-DEC-1998; 98KE-00062142.  
PR 08-MAR-1999; 99WO-US005028.  
PR 14-MAY-1999; 99US-00311832.  
PR 14-MAY-1999; 99WO-US010733.  
PR 25-AUG-1999; 99US-00380137.  
PR 25-AUG-1999; 99US-00380138.  
PR 25-AUG-1999; 99US-00380139.  
PR 25-AUG-1999; 99US-00380142.  
PR 15-SEP-1999; 99US-00397342.  
PR 18-OCT-1999; 99US-00403297.  
PR 12-NOV-1999; 99US-00423844.  
PR 30-DEC-1999; 99WO-US031274.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 02-JUN-2000; 2000US-00645264.  
PR 22-AUG-2000; 2000US-00644848.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 18-SEP-2000; 2000US-00664610.  
PR 18-SEP-2000; 2000US-00665350.  
PR 08-NOV-2000; 2000US-00703238.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 22-MAR-2001; 2001US-00816744.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 30-MAY-2001; 2001US-00870574.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 29-JUN-2001; 2001US-00869599.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-DEC-2001; 2001US-00006867.  
XX (GETH ) GENENTECH INC.  
XX  
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;  
XX  
DR WPI; 2003-447384/42.  
DR N-PSDB; ACD81547.  
XX  
XX New isolated antibody specifically binding a PRO polypeptide, useful for  
PT the preparation of a medicament for treating disorders with the aberrant  
PT expression or activity of the PRO polypeptide, such as tumor conditions  
PT and cancer.  
XX  
PS Claim 14; Fig 14; 223pp; English.  
XX  
CC The invention relates to an antibody that binds to a secreted or  
CC transmembrane protein designated PRO1446 appearing as ABO33941. The  
CC protein is one of 84 PRO polypeptides which (along with their encoding  
CC nucleic acids) are disclosed in the specification. The methods and  
CC compositions of the present invention are useful for the preparation of a  
CC medicament for the treatment of disorders associated with the aberrant  
CC expression or activity of the PRO polypeptide, such as tumour conditions  
CC and cancer. They can also be used to generate transgenic or knockout  
CC animals useful in the development and screening of therapeutically useful  
CC reagents. The PRO polypeptides and encoding nucleic acids can be used as  
CC molecular weight markers for protein electrophoresis, chromosome  
CC identification and tissue typing. The antibodies may be used in various  
CC diagnostic, competitive binding and/or immunoprecipitation assays. The  
CC present sequence represents a PRO polypeptide  
XX Sequence 234 AA;  
SQ

QY	Matches	6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Db	1 IETWFL 6									
	166 IETWFL 171									
RESULT 39										
ABR99384										
ID	ABR99384 standard; protein; 234 AA.									
XX										
AC	ABR99384;									
XX										
DT	18-SEP-2003 (first entry)									
XX										
DE	Human secreted polypeptide PRO1864, SEQ ID NO:66.									
XX										
KW	KW Human; PRO; secreted protein; transmembrane protein;									
KW	extracellular domain; tumour necrosis factor-alpha; TNF-alpha;									
KW	chondrocyte; proliferation; differentiation; cartilage disorder;									
KW	bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;									
KW	adrenal tumour; lung; colon; breast; prostate; kidney; cervix;									
KW	liver; drug screening; transgenic animal; genetic analysis;									
KW	antiarthritic; vulnary; gene therapy.									
OS	Homo sapiens.									
XX										
PN	US2003040063-A1.									
XX										
PD	27-FEB-2003.									
XX										
PF	26-JUN-2002; 2002US-00183006.									
XX										
PR	18-SEP-1997; 97US-0059263P.									
PR	18-SEP-1997; 97US-0059266P.									
PR	17-OCT-1997; 97US-0062250P.									
PR	21-OCT-1997; 97US-0063486P.									
PR	24-OCT-1997; 97US-0063120P.									
PR	24-OCT-1997; 97US-0063121P.									
PR	28-OCT-1997; 97US-0063540P.									
PR	28-OCT-1997; 97US-0063541P.									
PR	28-OCT-1997; 97US-0063544P.									
PR	29-OCT-1997; 97US-0063564P.									
PR	31-OCT-1997; 97US-0063870P.									
PR	31-OCT-1997; 97US-0064103P.									
PR	13-NOV-1997; 97US-0065311P.									
PR	21-NOV-1997; 97US-0066120P.									
PR	24-NOV-1997; 97US-0066466P.									
PR	24-NOV-1997; 97US-0066772P.									
PR	11-DEC-1997; 97US-0069335P.									
PR	12-DEC-1997; 97US-0069425P.									
PR	17-DEC-1997; 97US-0069870P.									
PR	18-DEC-1997; 97US-0068017P.									
PR	11-MAR-1998; 98US-0077450P.									
PR	11-MAR-1998; 98US-0077632P.									
PR	11-MAR-1998; 98US-0077649P.									
PR	20-MAR-1998; 98US-0078888P.									
PR	20-MAR-1998; 98US-0078939P.									
PR	27-MAR-1998; 98US-0079664P.									
PR	27-MAR-1998; 98US-0079786P.									
PR	31-MAR-1998; 98US-0080107P.									
PR	31-MAR-1998; 98US-0080194P.									
PR	01-APR-1998; 98US-0080327P.									
PR	01-APR-1998; 98US-0080333P.									
PR	08-APR-1998; 98US-0081049P.									
PR	08-APR-1998; 98US-0081070P.									
PR	09-APR-1998; 98US-0081195P.									
PR	15-APR-1998; 98US-0081838P.									
PR	21-APR-1998; 98US-0082568P.									
PR	21-APR-1998; 98US-0082569P.									
PR	22-APR-1998; 98US-0082704P.									
PR	22-APR-1998; 98US-0082797P.									



PR 28-APR-1998; 98US-00833222P.  
PR 29-APR-1998; 98US-00834495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083496P.  
PR 05-MAY-1998; 98US-0083559P.  
PR 06-MAY-1998; 98US-0083559P.  
PR 06-MAY-1998; 98US-0083559P.  
PR 07-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086332P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 05-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 18-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090439P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.

PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097352P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 03-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099755P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.

Query Match 66.7%; Score 6; DB 6; Length 234;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWFL 6  
Db 166 IETWFL 171

RESULT 40  
ABR98774  
ID ABR98774 standard; protein; 234 AA.  
XX  
AC ABR98774;  
XX  
DT 17-SEP-2003 (first entry)  
XX  
DE Human secreted polypeptide PRO1864, SEQ ID NO:66.  
XX  
KW Human; PRO; secreted protein; transmembrane protein;  
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;  
KW chondrocyte; proliferation; differentiation; cartilage disorder;  
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;  
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;  
KW liver; drug screening; transgenic animal; genetic analysis;  
KW antiarthritic; vulnery; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN US2003040064-A1.  
XX  
PD 27-FEB-2003.  
XX  
PF 26-JUN-2002; 2002US-00183008.  
XX  
PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 11-DEC-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078866P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 28-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 04-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.

```
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 25-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
```

```
Query Match 66.7%; Score 6; DB 6; Length 234;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IETWFL 6
Db 166 IETWFL 171
```

Search completed: August 31, 2006, 10:46:58  
Job time : 111.75 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:40:05 ; Search time 17.25 Seconds  
(without alignments)  
50.200 Million cell updates/sec

Title: DENGUE\_SEROTYPE2

Perfect score: 9

Sequence: 1 letwflrhp 9

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 1

Total number of hits satisfying chosen parameters: 283340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database :

PIR 80.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	55.6	115	2 A03862	hypothetical prote
2	5	55.6	132	2 S14077	IG kappa chain - A
3	5	55.6	137	2 A71308	hypothetical prote
4	5	55.6	168	2 B88102	protein W09G10.5 [
5	5	55.6	170	2 H69850	mutator Mutr prote
6	5	55.6	178	2 JQ1547	stripe disease-spe
7	5	55.6	196	2 F90534	transcription anti
8	5	55.6	208	2 T33341	hypothetical prote
9	5	55.6	210	2 A85018	probable copper-co
10	5	55.6	221	2 T35525	probable two compo
11	5	55.6	224	2 B87657	conserved hypoteth
12	5	55.6	225	2 C88939	protein C0584.8 [i
13	5	55.6	235	2 AF0656	conserved hypoteth
14	5	55.6	241	2 T27636	hypothetical prote
15	5	55.6	244	2 A12644	flagellar basal bo
16	5	55.6	244	2 A97427	flgF protein (U951
17	5	55.6	247	2 S75903	hypothetical prote
18	5	55.6	264	2 F85927	hypothetical prote
19	5	55.6	264	2 H65058	hypothetical prote
20	5	55.6	270	2 B64924	hypothetical prote
21	5	55.6	270	2 D85774	hypothetical prote
22	5	55.6	270	2 H90925	hypothetical prote
23	5	55.6	296	2 E91027	hypothetical prote
24	5	55.6	296	2 F85871	hypothetical prote
25	5	55.6	296	2 G65002	hypothetical prote
26	5	55.6	300	2 T32681	hypothetical prote
27	5	55.6	304	2 D64122	hypothetical prote
28	5	55.6	312	2 F82044	GGDEF family prote
29	5	55.6	313	2 AH0966	conserved hypoteth

30	5	55.6	317	2 S48036	hypothetical prote
31	5	55.6	321	2 AE1068	probable membrane
32	5	55.6	343	2 H64491	hypothetical prote
33	5	55.6	347	2 A46567	tetracycline resis
34	5	55.6	358	2 T34382	hypothetical prote
35	5	55.6	368	2 T06460	anthranilate phosp
36	5	55.6	375	1 B64593	2-oxoacid-ferredox
37	5	55.6	375	2 G71919	chain of 2-oxoglut
38	5	55.6	378	2 H96773	hypothetical prote
39	5	55.6	396	2 F58168	growth factor arg3
40	5	55.6	415	2 AC3235	nitroloxiacetate
41	5	55.6	415	2 B86434	protein T17H7.13 [
42	5	55.6	442	1 B69997	nitroloxiacetate
43	5	55.6	445	2 T38027	MLN 64 protein - h
44	5	55.6	445	2 E98303	hypothetical 49.3K
45	5	55.6	450	2 AH2979	nitroloxiacetate
46	5	55.6	456	2 T06589	3-methyl-2-oxobuta
47	5	55.6	473	2 JC4313	keratin 16, type I
48	5	55.6	484	2 E91082	hypothetical prote
49	5	55.6	601	2 F64116	endopeptidase La h
50	5	55.6	639	1 F69280	iron (II) transpor
51	5	55.6	720	2 S75935	hypothetical prote
52	5	55.6	722	2 T21521	hypothetical prote
53	5	55.6	759	2 B83474	probable type II s
54	5	55.6	783	2 B86254	hypothetical prote
55	5	55.6	818	2 F82173	collagenase VC1650
56	5	55.6	905	2 I49499	alpha N-catenin I
57	5	55.6	906	2 A43000	alpha N-catenin -
58	5	55.6	945	1 A45011	alpha-catenin 2 -
59	5	55.6	946	2 T31488	hypothetical prote
60	5	55.6	953	2 I49500	alpha N-catenin II
61	5	55.6	1011	1 A45598	H+-exporting ATPas
62	5	55.6	1155	2 B96761	probable protein k
63	5	55.6	1332	2 F69732	PBSX prophage ORF
64	5	55.6	1841	2 T38091	cell division cont
65	4	44.4	28	2 FX0033	cytochrome P450 te
66	4	44.4	44	1 WMV2K4	K4 protein - vacci
67	4	44.4	45	2 JH0208	hypothetical 5.2K
68	4	44.4	50	2 A69055	hypothetical prote
69	4	44.4	61	2 F96005	hypothetical prote
70	4	44.4	63	2 T15583	hypothetical prote
71	4	44.4	65	2 D87622	hypothetical prote
72	4	44.4	67	2 S08458	hypothetical prote
73	4	44.4	69	2 A71084	hypothetical prote
74	4	44.4	72	2 T13355	probable ribosomal
75	4	44.4	73	2 A90885	hypothetical prote
76	4	44.4	73	2 F85723	hypothetical prote
77	4	44.4	78	2 D29653	hypothetical prote
78	4	44.4	80	2 AF2836	hypothetical prote
79	4	44.4	81	2 B84095	hypothetical prote
80	4	44.4	82	2 H64896	probable membrane
81	4	44.4	87	2 C49917	probable pyruvate,
82	4	44.4	88	2 AF0549	conserved hypoteth
83	4	44.4	89	2 T42967	hypothetical prote
84	4	44.4	89	2 T50245	hypothetical prote
85	4	44.4	89	2 A50100	hypothetical prote
86	4	44.4	94	2 D64446	hypothetical prote
87	4	44.4	94	2 T29563	hypothetical prote
88	4	44.4	95	2 T18160	hypothetical prote
89	4	44.4	97	2 A10538	hypothetical prote
90	4	44.4	97	2 D95328	hypothetical prote
91	4	44.4	98	2 I49562	alpha-1 type III c
92	4	44.4	98	2 T17924	hypothetical prote
93	4	44.4	99	2 B90063	hypothetical prote
94	4	44.4	101	2 S37929	hypothetical prote
95	4	44.4	102	2 A12711	hypothetical prote
96	4	44.4	102	2 G97493	hypothetical prote
97	4	44.4	103	2 S64330	probable membrane
98	4	44.4	105	2 A72735	hypothetical prote
99	4	44.4	105	2 G72572	hypothetical prote
100	4	44.4	105	2 H81747	conserved hypoteth
101	4	44.4	106	2 T12684	hypothetical prote
102	4	44.4	106	2 S51046	hypothetical prote

103 4 44.4 106 2 E90062  
104 4 44.4 109 2 T29627  
105 4 44.4 110 2 S64948  
106 4 44.4 112 2 T02744  
107 4 44.4 112 2 AD3596  
108 4 44.4 114 2 D84852  
109 4 44.4 114 2 D72665  
110 4 44.4 115 2 G72568  
111 4 44.4 115 2 F72779  
112 4 44.4 116 2 H95414  
113 4 44.4 118 2 G59121  
114 4 44.4 120 2 C86882  
115 4 44.4 121 2 S76514  
116 4 44.4 122 2 D72756  
117 4 44.4 123 2 H90236  
118 4 44.4 123 2 S09822  
119 4 44.4 123 2 AH2707  
120 4 44.4 124 2 S03521  
121 4 44.4 124 2 C75359  
122 4 44.4 125 2 F82834  
123 4 44.4 127 2 C75315  
124 4 44.4 128 2 S76468  
125 4 44.4 128 2 AF2143  
126 4 44.4 129 2 A75346  
127 4 44.4 130 2 A41911  
128 4 44.4 132 2 T50389  
129 4 44.4 133 2 B84087  
130 4 44.4 133 2 B44370  
131 4 44.4 134 2 F85589  
132 4 44.4 134 2 D90739  
133 4 44.4 134 2 F64817  
134 4 44.4 134 2 C84385  
135 4 44.4 135 2 S31682  
136 4 44.4 136 2 T18052  
137 4 44.4 136 2 H72633  
138 4 44.4 137 2 A95229  
139 4 44.4 137 2 E98093  
140 4 44.4 137 2 F29380  
141 4 44.4 137 2 S40760  
142 4 44.4 138 2 T13557  
143 4 44.4 140 2 S42424  
144 4 44.4 141 1 HALZC  
145 4 44.4 141 2 J0624  
146 4 44.4 141 2 T46427  
147 4 44.4 142 2 F82239  
148 4 44.4 143 2 A10838  
149 4 44.4 145 2 AC2452  
150 4 44.4 146 2 T35484

## ALIGNMENTS

RESULT 1  
A03862  
hypothetical protein E-115 - human adenovirus 2  
C:Species: Mastadenovirus h2 (human adenovirus 2)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C:Accession: A03862  
R:Gingeras, T.R.; Sciaky, D.; Gelinias, R.E.; Bing-Dong, J.; Yen, C.E.; Kelly, M.M.; Bull  
J. Biol. Chem. 257, 13475-13491, 1982  
A:Title: Nucleotide sequences from the adenovirus-2 genome.  
A:Reference number: A92351; MUID:83056843; PMID:7142161  
A:Accession: A03862  
A:Molecule type: DNA  
A:Residues: 1-115 <GIN>  
A:Cross-references: UNIPROT:P03290; UNIPARC:UPI00001392B5

Query Match 55.6%; Score 5; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 ETWFL 6

Db 3 ETWFL 7  
RESULT 2  
IG kappa chain - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999  
C:Accession: S14077  
R:Schwager, J.; Buerckert, N.; Schwager, M.; Wilson, M.  
EMBO J. 10, 505-511, 1991  
A:Title: Evolution of immunoglobulin light chain genes: analysis of Xenopus Igl isotypes  
A:Reference number: S14076; MUID:91160503; PMID:1705882  
A:Accession: S14077  
A:Molecule type: mRNA  
A:Residues: 1-132 <SCH>  
A:Cross-references: UNIPARC:UPI000017698D  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
Query Match 55.6%; Score 5; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
3 TWFLR 7  
48 TWFLR 52  
Db  
RESULT 3  
A71308  
hypothetical protein TP0563 - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: A71308  
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDor  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770; PMID:9665876  
A:Accession: A71308  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-137 <COL>  
A:Cross-references: UNIPROT:O83574; UNIPARC:UPI00000C0A8E; GB:AE001232; GB:AE000520; NID:  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0563  
C:Superfamily: syphilis spirochete hypothetical protein TP0563  
Query Match 55.6%; Score 5; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 IETWF 5  
122 IETWF 126  
Db  
RESULT 4  
B88102  
protein W09G10.5 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: B88102  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_eleg  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A;Accession: B88102  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-168 <STO>  
A;Cross-references: UNIPROT:O16641; UNIPARC:UPI0000075172; GB:chr\_II; PIDN:AB66113.1; PIDN:AB66113.1  
C;Genetics:  
A;Gene: W09G10.5  
A;Map position: 2  
C;Superfamily: Caenorhabditis elegans hypothetical protein C31G12.2

Query Match 55.6%; Score 5; DB 2; Length 168;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
|||||  
Db 30 TWFLR 34

RESULT 5  
H69850  
mutator MutT protein homolog yjhb - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: H69850  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea  
C; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Etrington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Segkiguchi, J.; Sekowska, A.; Seroh  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: H69850  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-170 <KUN>  
A;Cross-references: UNIPROT:O34488; UNIPARC:UPI0000060253; GB:Z59110; GB:AL009126; NID:9  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: yjhb

Query Match 55.6%; Score 5; DB 2; Length 170;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 159 FLRHP 163

RESULT 6  
JQ1547  
stripe disease-specific protein - rice stripe virus (isolate T)  
C;Species: rice stripe virus  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: JQ1547  
R;Zhu, Y.; Hayakawa, T.; Toriyama, S.  
J. Gen. Virol. 73, 1309-1312, 1992  
A;Title: Complete nucleotide sequence of RNA 4 of rice stripe virus isolate T, and comp  
A;Reference number: JQ1547; MUID:92268894; PMID:1588328  
A;Accession: JQ1547  
A;Molecule type: genomic RNA  
A;Residues: 1-178 <ZHU>  
A;Cross-references: UNIPROT:Q00844; UNIPARC:UPI0000002ECF; GB:D10979; DBV:D01164; NID:9

Query Match 55.6%; Score 5; DB 2; Length 208;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
|||||  
Db 159 TWFLR 163

RESULT 9  
A85018  
probable copper-containing glycoprotein [imported] - Arabidopsis thaliana

C;Superfamily: maize stripe virus major noncapsid protein

Query Match 55.6%; Score 5; DB 2; Length 178;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 123 FLRHP 127

RESULT 7  
F90534  
transcription antitermination protein [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C;Species: Mycoplasma pulmonis  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C;Accession: F90534  
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm  
A;Reference number: A99512; MUID:21267165; PMID:11353084  
A;Accession: F90534  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-196 <KUR>  
A;Cross-references: UNIPROT:Q98R28; UNIPARC:UPI00000D4589; GB:AL445566; PID:gl4089595; P  
A;Experimental source: strain UAB CTIP  
C;Genetics:  
A;Gene: MYPU 1820  
A;Genetic code: SGC3

Query Match 55.6%; Score 5; DB 2; Length 196;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETWFL 6  
|||||  
Db 82 ETWFL 86

RESULT 8  
T33341  
hypothetical protein K07D4.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 31-Dec-2004  
C;Accession: T33341  
R;Henkhaus, J.; Wohldmann, P.  
A;Description: The sequence of C. elegans cosmid K07D4.  
submitted to the EMBL Data Library, July 1998  
A;Reference number: Z21327  
A;Accession: T33341  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-208 <HEN>  
A;Cross-references: UNIPROT:O76574; UNIPARC:UPI000007CD7A; EMBL:AF077534; PIDN:AAC26289.3  
A;Experimental source: strain Bristol N2; clone K07D4  
C;Genetics:  
A;Gene: CESP:K07D4.5  
A;Map position: 2  
A;Introns: 25/3; 68/1; 127/1; 160/2

Query Match 55.6%; Score 5; DB 2; Length 208;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
|||||  
Db 159 TWFLR 163

RESULT 9  
A85018  
probable copper-containing glycoprotein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: A85018  
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999  
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A:Reference number: A85001; MUID:20083488; PMID:10617198  
 A:Accession: A85018  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-210 <STO>  
 A:Cross-references: UNIPROT:Q9M135; UNIPARC:UPI000009E60F; GB:NC\_001368; NID:g7267635; E  
 C:Genetics:  
 A:Gene: AT4g01380  
 A:Map position: 4

Query Match 55.6%; Score 5; DB 2; Length 210;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWF 5  
 |||||  
 DB 18 IETWF 22

RESULT 10  
 T35525  
 probable two component response regulator - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
 C:Accession: T35525  
 R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, March 1999  
 A:Reference number: Z21581  
 A:Accession: T35525  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-221 <SEE>  
 A:Cross-references: UNIPROT:Q9X802; UNIPARC:UPI00000DAF6B; EMBL:AL049497; PIDN:CAB39870.  
 C:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCODB:SC6G10.16  
 C:Superfamily: ompR protein; response regulator homology

Query Match 55.6%; Score 5; DB 2; Length 221;  
 Best Local Similarity 100.0%; Pred. No. 54;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 |||||  
 DB 163 FLRHP 167

RESULT 11  
 B87657  
 conserved hypothetical protein CC3292 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C:Accession: B87657  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: B87657  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-224 <STO>  
 A:Cross-references: UNIPROT:Q9A3B2; UNIPARC:UPI00000C7A0F; GB:AE005673; NID:g13424986; E  
 C:Genetics:  
 A:Gene: CC3292

Query Match 55.6%; Score 5; DB 2; Length 224;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
 |||||  
 DB 97 WFLRH 101

RESULT 12  
 C88939  
 protein C05B4.8 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 C:Accession: C88939  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_eleg  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: C88939  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-225 <STO>  
 A:Cross-references: UNIPROT:O17356; UNIPARC:UPI0000082F2F; GB:chr\_V; PIDN:AAB71277.1; PII  
 A:Note: similar to C. elegans TC3 transposase (SP:34257)  
 C:Genetics:  
 A:Gene: C05B4.8  
 A:Map position: 5  
 C:Superfamily: Caenorhabditis transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 |||||  
 DB 184 FLRHP 188

RESULT 13  
 AF0656  
 conserved hypothetical protein STY1354 [imported] - Salmonella enterica subsp. enterica s  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AF0656  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AF0656  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-235 <PAR>  
 A:Cross-references: UNIPARC:UPI0000059P2E; GB:AL513382; PIDN:CAD01623.1; PID:g16502477; C  
 C:Genetics:  
 A:Gene: STY1354

Query Match 55.6%; Score 5; DB 2; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
 |||||  
 DB 65 WFLRH 69

RESULT 14



T27636  
 hypothetical protein ZC64.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T27636  
 R:Bentley, D.  
 submitted to the EMBL Data Library, October 1995  
 A:Description: The sequence of *C. elegans* cosmid ZC64.  
 A:Reference number: Z20397  
 A:Accession: T27636  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-241 <BEN>  
 A:Cross-references: UNIPROT:Q23379; UNIPARC:UPI0000081BA9; EMBL:U39740; PIDN:AAA80427.1;  
 C:Genetics:  
 A:Gene: CESP:ZC64.1  
 C:Superfamily: *Caenorhabditis* transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 Db 204 FLRHP 208

RESULT 15  
 A12644  
 flagellar basal body rod protein [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)  
 C:Species: *Agrobacterium tumefaciens*  
 C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 05-Oct-2004  
 C:Accession: A12644  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, S.; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.; E.W.  
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: A12644  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-244 <KUR>  
 A:Cross-references: UNIPROT:Q34170; UNIPARC:UPI00000D1464; GB:AE008688; PIDN:AAL41575.1;  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: flgF  
 A:Map position: circular chromosome  
 C:Superfamily: rod protein flgF

Query Match 55.6%; Score 5; DB 2; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 Db 168 FLRHP 172

RESULT 16  
 A97427  
 flgF protein (U95165) [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)  
 C:Species: *Agrobacterium tumefaciens*  
 C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 05-Oct-2004  
 C:Accession: A97427  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* C58  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: A97427

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-244 <KUR>  
 A:Cross-references: UNIPROT:Q34170; UNIPARC:UPI00000D1464; GB:AE007869; PIDN:AAK86370.1;  
 C:Genetics:  
 A:Gene: AGR\_C\_982  
 A:Map position: circular chromosome  
 C:Superfamily: rod protein flgF

Query Match 55.6%; Score 5; DB 2; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 Db 168 FLRHP 172

RESULT 17  
 S75903  
 hypothetical protein - *Synechocystis* sp. (strain PCC 6803)  
 C:Species: *Synechocystis* sp.  
 A:Variety: PCC 6803  
 C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
 C:Accession: S75903  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S75903  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-247 <KAN>  
 A:Cross-references: UNIPROT:P74268; UNIPARC:UPI00001290CE; EMBL:D90913; GB:AB001339; NID A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 55.6%; Score 5; DB 2; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 Db 209 FLRHP 213

RESULT 18  
 F85927  
 hypothetical protein Z4084 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL9:  
 C:Species: *Escherichia coli*  
 C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: F85927  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: F85927  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-264 <STO>  
 A:Cross-references: UNIPARC:UPI00001658F2; GB:AE005174; NID:gl2517235; PIDN:AAG57882.1;  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: Z4084

Query Match 55.6%; Score 5; DB 2; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWF 5  
 Db 1 IETWF 5

Db 99 IETWF 103

## RESULT 19

H65058

hypothetical protein b2772 - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C&gt;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002

C:Accession: H65058

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: H65058

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-264 &lt;BLAT&gt;

A:Cross-references: UNIPARC:UPI0000168160; GB:AE000360; GB:U00096; NID:g2367157; PIDN:AE

A:Experimental source: strain K-12, substrain MG1655

Query Match 55.6%; Score 5; DB 2; Length 264;

Best Local Similarity 100.0%; Pred. No. 63;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWF 5

Db 99 IETWF 103

## RESULT 20

E64924

hypothetical protein b1669 - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C&gt;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004

C:Accession: E64924

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: E64924

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-270 &lt;BLAT&gt;

A:Cross-references: UNIPROT:P77147; UNIPARC:UPI000013A9BC; GB:AE000262; GB:U00096; NID:9

A:Experimental source: strain K-12, substrain MG1655

C:Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8

Db 20 WFLRH 24

## RESULT 21

D85774

hypothetical protein Z2696 [imported] - Escherichia coli (strain O157:H7, substrain EDL9

C:Species: Escherichia coli

C&gt;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C:Accession: D85774

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D85774

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-270 &lt;STO&gt;

A:Cross-references: UNIPROT:Q8X618; UNIPARC:UPI00000D0C14; GB:AE005174; NID:g12515668; P1  
 A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z2696

C:Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8

Db 20 WFLRH 24

## RESULT 22

H90925

hypothetical protein ECS2376 [imported] - Escherichia coli (strain O157:H7, substrain RIN

C:Species: Escherichia coli

C&gt;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004

C:Accession: H90925

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H90925

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-270 &lt;HAY&gt;

A:Cross-references: UNIPROT:Q8X618; UNIPARC:UPI0000D0C14; GB:BA000007; PIDN:BA035799.1;

A:Experimental source: strain O157:H7, substrain RIND 0509952

C:Genetics:

A:Gene: ECS2376

C:Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8

Db 20 WFLRH 24

## RESULT 23

E91027

hypothetical protein ECS189 [imported] - Escherichia coli (strain O157:H7, substrain RIN

C:Species: Escherichia coli

C&gt;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 31-Dec-2004

C:Accession: E91027

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: E91027

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-296 &lt;HAY&gt;

A:Cross-references: UNIPROT:Q8XCT0; UNIPARC:UPI0000D0433; GB:BA000007; PIDN:BA036612.1;

A:Experimental source: strain O157:H7, substrain RIND 0509952

C:Genetics:

A:Gene: ECS189

C:Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 296;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9

Db 17 FLRHP 21

```

RESULT 24
F85871
hypothetical protein yfci [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
C;Accession: F85871
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85871
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-296 <STO>
A;Cross-references: UNIPROT:Q8XCT0; UNIPARC:UPI000000D0433; GB:AE005174; NID:gl2516661; F
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
C;Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 17 FLRHP 21

RESULT 25
G65002
hypothetical protein b2305 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 31-Dec-2004
C;Accession: G65002
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: G65002
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-296 <BLAT>
A;Cross-references: UNIPROT:P77768; UNIPARC:UPI00000047C88; GB:AE000319; GB:U00096; NID:g
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 17 FLRHP 21

RESULT 26
T32681
hypothetical protein K07C6.14 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32681
R;Wagner-McPherson, C.; Gillam, B.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid K07C6.
A;Reference number: Z21209
A;Accession: T32681
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

```

```

A;Residues: 1-300 <WAG>
A;Cross-references: UNIPROT:O44645; UNIPARC:UPI00000079677; EMBL:AF039049; PIDN:AAB94256.
A;Experimental source: strain Bristol N2; clone K07C6
C;Genetics:
A;Gene: CESP:K07C6.14
A;Map position: 5
C;Superfamily: Caenorhabditis transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 231 FLRHP 235

RESULT 27
D64122
hypothetical protein H11424 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: D64122
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: D64122
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-304 <TIGR>
A;Cross-references: UNIPROT:P45198; UNIPARC:UPI0000013AB2; GB:U32821; GB:L42023; NID:g15

Query Match 55.6%; Score 5; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 60 FLRHP 64

RESULT 28
F82044
GGDEF family protein VC2697 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: F82044
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: F82044
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <HEI>
A;Cross-references: UNIPROT:Q9KNN4; UNIPARC:UPI000000C3389; GB:AE004335; GB:AE003852; NID
A;Experimental source: serogroup O1, strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2697
A;Map position: 1

Query Match 55.6%; Score 5; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWF 5
|||||

```



Db 313 TWFLR 317

RESULT 34  
T34382  
hypothetical protein T25G12.9 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T34382  
R;Du, Z.  
submitted to the EMBL Data Library, December 1995  
A;Description: The sequence of *C. elegans* cosmid T25G12.  
A;Reference number: Z21515  
A;Accession: T34382  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-358 <DUZ>  
A;Cross-references: UNIPROT:Q22789; UNIPARC:UPI0000081C10; EMBL:U43283; PIDN:NAAC69023.1;  
A;Experimental source: strain Bristol N2; clone T25G12  
C;Genetics:  
A;Gene: CESP:T25G12.9  
A;Map position: X  
C;Superfamily: *Caenorhabditis* transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 358;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
|||||  
Db 231 FLRHP 235

RESULT 35  
T06460  
anthranilate phosphoribosyltransferase (EC 2.4.2.18) - garden pea (fragment)  
N;Alternate names: phosphoribosylanthranilate transferase  
C;Species: *Pisum sativum* (garden pea)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: T06460  
R;Sato, N.; Kazuno, A.; Ohta, N.; Ohehima, K.  
submitted to the EMBL Data Library, June 1996  
A;Description: Isolation of a pea cDNA for phosphoribosylanthranilate transferase.  
A;Reference number: Z15694  
A;Accession: T06460  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-368 <SAT>  
A;Cross-references: UNIPROT:Q43085; UNIPARC:UPI000000A9D1C; EMBL:D86180; PIDN:BAAL3032.1  
A;Experimental source: var. Alaska  
C;Genetics:  
A;Gene: PAT1  
C;Keywords: glycosyltransferase; pentosyltransferase

Query Match 55.6%; Score 5; DB 2; Length 368;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
|||||  
Db 339 FLRHP 343

RESULT 36  
E64593  
2-oxoacid-ferredoxin oxidoreductase (EC 1.2.7.-) alpha chain - *Helicobacter pylori* (strain N)  
N;Alternate names: 2-oxoacid:ferredoxin oxidoreductase (CoA-acetylating)  
C;Species: *Helicobacter pylori*  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: E64593  
R;Tomb, J.F.; White, O.; Kervilavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.; A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: E64593  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-375 <TOM>  
A;Cross-references: UNIPROT:O25311; UNIPARC:UPI00000D30AB; GB:AE000572; GB:AE000511; NID  
C;Superfamily: *Helicobacter pylori* 2-oxoacid ferredoxin oxidoreductase; 2-oxoacid ferredoxin oxidoreductase  
C;Keywords: oxidoreductase  
F;5-186/Domain: 2-oxoacid ferredoxin oxidoreductase homology <FEO>

Query Match 55.6%; Score 5; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
|||||  
Db 124 FLRHP 128

RESULT 37  
G71919  
chain of 2-oxoglutarate oxidoreductase - *Helicobacter pylori* (strain J99)  
C;Species: *Helicobacter pylori*  
A;Variety: strain J99  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: G71919  
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*.  
A;Reference number: A71800; MUID:99120557; PMID:9923682  
A;Accession: G71919  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-375 <ARN>  
A;Cross-references: UNIPROT:Q9ZLPI; UNIPARC:UPI00000D364E; GB:AE001486; GB:AE001439; NID  
A;Experimental source: strain J99  
C;Genetics:  
A;Gene: OorA  
C;Superfamily: *Helicobacter pylori* 2-oxoacid ferredoxin oxidoreductase; 2-oxoacid ferredoxin oxidoreductase  
F;5-186/Domain: 2-oxoacid ferredoxin oxidoreductase homology <FEO>

Query Match 55.6%; Score 5; DB 2; Length 375;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
|||||  
Db 124 FLRHP 128

RESULT 38  
H96773  
hypothetical protein F1M20.17 [imported] - *Arabidopsis thaliana*  
C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: H96773  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: H96773  
A;Status: preliminary

```

A:Molecule type: DNA
A:Residues: 1-378 <STO>
A:Cross-references: UNIPROT:Q9CAG5; UNIPARC:UPI00000A3E86; GB:AE005173; NID:g6539251; PID:
C:Genetics:
A:Gene: FIM20.17
A:Map position: 1
C:Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match      55.6%; Score 5; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 FLRHP 9
Db      372 FLRHP 376

RESULT 39
I58168
Growth factor arg3.1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: I58168; I59386
R:Lyford, G.L.; Yamagata, K.; Kaufmann, W.E.; Barnes, C.A.; Sanders, L.K.; Copeland, N.G.
Neuron 14, 433-445, 1995
A:Title: Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-
A:Reference number: I58168; MUID:95161073; PMID:7857651
A:Accession: I58168
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-396 <RES>
A:Cross-references: UNIPROT:Q62743; UNIPARC:UPI00000E5C7E; EMBL:U19866; NID:g644828; PID
R:Link, W.; Konietzko, U.; Kauselmann, G.; Krug, M.; Schwanke, B.; Frey, U.; Kuhl, D.
Proc. Natl. Acad. Sci. U.S.A. 92, 5734-5738, 1995
A:Title: Somatodendritic expression of an immediate early gene is regulated by synaptic
A:Reference number: I59386; MUID:95296386; PMID:7777577
A:Accession: I59386
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-208, 'V', 210-396 <RE2>
A:Cross-references: UNIPARC:UPI00000E79AC; EMBL:Z46925; NID:g854413; PIDN:CAA87033.1; PI
C:Genetics:
A:Gene: Arc
C:Superfamily: rat growth factor arg3.1

Query Match      55.8%; Score 5; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 FLRHP 9
Db      336 FLRHP 340

RESULT 40
AC3235
nitrotriacetate monooxygenase nrtA [imported] - Agrobacterium tumefaciens (strain C58,
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AC3235
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AC3235
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-415 <KUR>
A:Cross-references: UNIPROT:Q8U674; UNIPARC:UPI00000D276A; GB:AE008690; PIDN:AAL46297.1;

```

```

A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: nrtA
A:Genome: plasmid
C:Superfamily: nitrotriacetate monooxygenase

```

```

Query Match      55.6%; Score 5; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ETWFL 6
Db      348 ETWFL 352

```

```

Search completed: August 31, 2006, 10:48:04
Job time : 18.25 secs

```

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:29:54 ; Search time 139.25 Seconds  
(without alignments)  
59.786 Million cell updates/sec

Title: DENGUE\_SEROTYPE2

Perfect score: 9

Sequence: 1 ietwflrhp 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2849466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database :

1: uniprot\_7.2.2\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	77.8	266	2	Q5FQ40 GLUCONOBACT
2	6	66.7	234	1	Q95772 HOMO SAPIEN
3	6	66.7	235	1	Q9DC13 MUS MUSCULU
4	6	66.7	235	2	Q3U852 MUS MUSCULU
5	6	66.7	235	2	Q3U8Q7 MOUSE
6	6	66.7	235	2	Q5U205 RAT
7	6	66.7	405	2	Q4IRY2 GIBZEB
8	6	66.7	1079	1	GP113 HUMAN
9	6	66.7	1079	2	Q53TA5 HUMAN
10	6	66.7	1093	2	Q8X0R0 NEUCR
11	6	66.7	1117	2	Q7S795 NEUCR
12	5	55.6	29	2	Q6CGE3 YARULI
13	5	55.6	79	2	Q8A2T7 BACTN
14	5	55.6	77	2	Q71136 LACDCL
15	5	55.6	80	2	Q46AP0 METAC
16	5	55.6	80	2	Q8TIM1 METAC
17	5	55.6	87	2	Q5QCQ7 CENAS
18	5	55.6	95	2	Q9MIS5 PELLE
19	5	55.6	101	2	Q3B3X5 PELDICTYON
20	5	55.6	109	2	Q82LF3 STREPTOMYCE
21	5	55.6	115	1	Y115 ADE02
22	5	55.6	115	2	Q2KSZ2 ADE05
23	5	55.6	118	2	Q3WDX0 GACTO
24	5	55.6	118	2	Q4QKE8 HAEB18
25	5	55.6	122	2	Q5P5F8 AZOSE
26	5	55.6	124	2	Q47S60 THEFY
27	5	55.6	127	1	CRCB ERWCT
28	5	55.6	128	2	Q90Z26 XENTR
29	5	55.6	130	2	Q2SYI4 BURTH
30	5	55.6	130	2	Q4HDP5 CAMCO
31	5	55.6	132	2	Q6T1W1 ANETH

32	5	55.6	133	2	Q62HE5 BURWA	Q62he5 burkholderi
33	5	55.6	133	2	Q63R57 BURPS	Q63r57 burkholderi
34	5	55.6	134	2	Q411L0 KINRA	Q411l0 kinococcus
35	5	55.6	136	2	Q977K8 SCREN	Q977k8 uncultured
36	5	55.6	136	2	Q82HV8 STRAW	Q82hv8 streptomyce
37	5	55.6	137	2	Q83574 TREPA	Q83574 treponema p
38	5	55.6	137	2	Q73P38 TREDE	Q73p38 treponema d
39	5	55.6	139	2	Q6IGY3 DROME	Q6igy3 drephonilla
40	5	55.6	139	2	Q4NBA0 9MICC	Q4nba0 arthrobacte
41	5	55.6	140	2	Q6Z0F6 ORVSA	Q6z0f6 oryza sativ
42	5	55.6	141	2	Q3Y4I1 9BACT	Q3y4i1 uncultured
43	5	55.6	141	2	Q67WA3 SYMPH	Q67wa3 symbiobacte
44	5	55.6	145	2	Q9RIK1 STREY	Q9rik1 streptococc
45	5	55.6	151	2	Q417C9 KINRA	Q417c9 kinococcus
46	5	55.6	157	2	Q849K6 STRVN	Q849k6 streptomyce
47	5	55.6	158	2	Q2PS78 9BACT	Q2ps78 uncultured
48	5	55.6	158	2	Q5KYB0 GEOKA	Q5kyb0 geobacillus
49	5	55.6	163	2	Q7NJ45 GLOVI	Q7nj45 gloeobacter
50	5	55.6	166	2	Q2KC79 RHET	Q2kc79 rhizobium e
51	5	55.6	168	2	Q16641 CAEEL	Q16641 caenorhabdi
52	5	55.6	170	2	Q34488 BACSM	Q34488 bacillus su
53	5	55.6	171	2	Q2S721 9GAMM	Q2s721 haella che
54	5	55.6	171	2	Q6X8U7 9BACT	Q6x8u7 uncultured
55	5	55.6	171	2	Q6X8U8 9BACT	Q6x8u8 uncultured
56	5	55.6	171	2	Q8ZJZ4 SALTY	Q8zjz4 salmoneilla
57	5	55.6	171	2	Q9EWQ7 STRCO	Q9ewq7 streptomyce
58	5	55.6	172	2	Q6X8Q2 9BACT	Q6x8q2 uncultured
59	5	55.6	172	2	Q6X8Q6 9BACT	Q6x8q6 uncultured
60	5	55.6	172	2	Q6X8Q9 9BACT	Q6x8q9 uncultured
61	5	55.6	172	2	Q6X8S4 9BACT	Q6x8s4 uncultured
62	5	55.6	172	2	Q6X8S5 9BACT	Q6x8s5 uncultured
63	5	55.6	172	2	Q6X8T2 9BACT	Q6x8t2 uncultured
64	5	55.6	172	2	Q6X8V0 9BACT	Q6x8v0 uncultured
65	5	55.6	172	2	Q768U7 9BACT	Q768u7 uncultured
66	5	55.6	172	2	Q768V6 9BACT	Q768v6 uncultured
67	5	55.6	172	2	Q8KPU7 9BACT	Q8kpu7 uncultured
68	5	55.6	172	2	Q8KPK3 9BACT	Q8kpk3 uncultured
69	5	55.6	175	2	Q98II9 RHILIO	Q98ii9 rhizobium l
70	5	55.6	176	2	Q4MIV5 BACCE	Q4miv5 bacillus ce
71	5	55.6	176	2	Q737H4 BACC1	Q737h4 bacillus ce
72	5	55.6	178	1	VNCA_RSVN	Q01209 rice stripe
73	5	55.6	178	1	VNCA_RSVT	Q00844 rice stripe
74	5	55.6	178	2	Q4TUA0 9VIRU	Q4tua0 rice stripe
75	5	55.6	178	2	Q52P74 9VIRU	Q52p74 rice stripe
76	5	55.6	178	2	Q52R45 9VIRU	Q52r45 rice stripe
77	5	55.6	178	2	Q52R48 9VIRU	Q52r48 rice stripe
78	5	55.6	178	2	Q52R49 9VIRU	Q52r49 rice stripe
79	5	55.6	178	2	Q6EWP1 9VIRU	Q6ewp1 rice stripe
80	5	55.6	178	2	Q705B3 9VIRU	Q705b3 rice stripe
81	5	55.6	178	2	Q705B5 9VIRU	Q705b5 rice stripe
82	5	55.6	178	2	Q705B6 9VIRU	Q705b6 rice stripe
83	5	55.6	178	2	Q71TU0 9VIRU	Q71tu0 rice stripe
84	5	55.6	178	2	Q7TL11 9VIRU	Q7tl11 rice stripe
85	5	55.6	178	2	Q7TL12 9VIRU	Q7tl12 rice stripe
86	5	55.6	178	2	Q7TL13 9VIRU	Q7tl13 rice stripe
87	5	55.6	178	2	Q7TL15 9VIRU	Q7tl15 rice stripe
88	5	55.6	178	2	Q7TL16 9VIRU	Q7tl16 rice stripe
89	5	55.6	178	2	Q7TL17 9VIRU	Q7tl17 rice stripe
90	5	55.6	178	2	Q7TL18 9VIRU	Q7tl18 rice stripe
91	5	55.6	178	2	Q7TL19 9VIRU	Q7tl19 rice stripe
92	5	55.6	178	2	Q80A47 9VIRU	Q80a47 rice stripe
93	5	55.6	178	2	Q80A50 9VIRU	Q80a50 rice stripe
94	5	55.6	178	2	Q9ICC6 9VIRU	Q9icc6 rice stripe
95	5	55.6	178	2	Q9J0W5 9VIRU	Q9j0w5 rice stripe
96	5	55.6	178	2	Q9J0W6 9VIRU	Q9j0w6 rice stripe
97	5	55.6	178	2	Q9J0W7 9VIRU	Q9j0w7 rice stripe
98	5	55.6	178	2	Q5K015 9VIRU	Q5k015 rice stripe
99	5	55.6	178	2	Q5K018 9VIRU	Q5k018 rice stripe
100	5	55.6	178	2	Q5K003 9VIRU	Q5k003 rice stripe
101	5	55.6	178	2	Q5K011 9VIRU	Q5k011 rice stripe
102	5	55.6	178	2	Q5K013 9VIRU	Q5k013 rice stripe
103	5	55.6	178	2	Q4SBK7 TETNG	Q4sbk7 tetraodon n
104	5	55.6	183	2	Q5WJB7 BACSK	Q5wjb7 bacillus cl

```

105 5 55.6 190 2 Q6L1R8 PICTO
106 5 55.6 190 2 Q2UK03 ASPOR
107 5 55.6 191 2 Q3QVT3 RHOB
108 5 55.6 194 2 Q9B861 RHILLO
109 5 55.6 196 2 Q9R28 MYCPFU
110 5 55.6 198 2 Q3GC97 9FRM
111 5 55.6 198 2 Q8ZP58 SALTY
112 5 55.6 204 2 Q3F261 9BURK
113 5 55.6 204 2 Q3G826 9DELT
114 5 55.6 207 2 Q8NN53 CORGL
115 5 55.6 209 2 Q21P50 9DELT
116 5 55.6 210 2 Q9M135 ARATH
117 5 55.6 211 2 Q8R221 MOUSE
118 5 55.6 213 2 Q4B830 PSE14
119 5 55.6 213 2 Q6VEB2 PSEY1
120 5 55.6 215 2 Q5FJBI LACAC
121 5 55.6 218 2 Q5Z317 NOCFA
122 5 55.6 219 2 Q47QAL THEFY
123 5 55.6 220 2 Q3W859 9ACTO
124 5 55.6 220 2 Q4OV16 KINRA
125 5 55.6 221 2 Q9X802 STRCO
126 5 55.6 224 2 Q6F256 MESFL
127 5 55.6 224 2 Q9A3B2 CAUCR
128 5 55.6 225 2 Q331V9 METHU
129 5 55.6 225 2 Q4H823 9DEIO
130 5 55.6 231 2 Q4BKG1 CROWT
131 5 55.6 232 2 Q85SH1 ORISA
132 5 55.6 232 2 Q2T527 BURTH
133 5 55.6 235 2 Q57NW5 SALCH
134 5 55.6 235 2 Q5PIM7 SALPA
135 5 55.6 235 2 Q8Z7C5 SALTU
136 5 55.6 237 2 Q44AD1 SOLUS
137 5 55.6 238 2 Q7ND55 GLOVI
138 5 55.6 239 2 Q7QWJ0 GIALLA
139 5 55.6 244 2 Q16195 HUMAN
140 5 55.6 244 2 Q34170 9RHIZ
141 5 55.6 244 2 Q7D183 AGRIT5
142 5 55.6 245 2 Q93SE2 ECOLI
143 5 55.6 245 2 Q8XX11 SYN2P
144 5 55.6 245 2 Q8ZGS8 STRAW
145 5 55.6 246 2 Q4UD31 THEAN
146 5 55.6 247 1 DDPX SYNY3
147 5 55.6 248 2 Q5W912 XENTR
148 5 55.6 249 2 Q8ZF47 STRAW
149 5 55.6 253 2 Q53F67 HUMAN
150 5 55.6 254 2 Q4N0T1 THEPRA

```

## ALIGNMENTS

```

RESULT 1
ID Q5FQA0_GLUOX PRELIMINARY; PRT; 266 AA.
AC Q5FQA0;
DT 01-MAR-2005, integrated into UniProtKB/TrEMBL.
DE 07-FEB-2006, sequence version 1.
DE Putative hydrolase of the HAD superfamily.
GN OrderedLocusNames=GOX1706;
OS Gluconobacter oxydans (Gluconobacter suboxydans).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconobacter.
OX NCBI_TaxID=442;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=621H;
RX PubMed=15665824; DOI=10.1038/nbt1062;
RA Prust C., Hoffmeister M., Liesegang H., Wiezer A., Fricke W.F.,
RA Ehrenreich A., Gottschalk G., Deppenmeier U.;
RT "Complete genome sequence of the acetic acid bacterium Gluconobacter
RT oxydans."
RL Nat. Biotechnol. 23:195-200(2005).

```

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000009; AAW61446.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR013200; HAD_3.
DR InterPro; IPR006379; HAD_SF_IIB.
DR InterPro; IPR000150; Hypothet_cof.
DR Pfam; PF00702; Hydrolase; 1.
DR TIGRFAMs; TIGR00099; Cof-subfamily; 1.
DR TIGRFAMs; TIGR01484; HAD_SF_IIB; 1.
DR PROSITE; PS01229; COF_2; UNKNOWN_1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 266 AA; 28559 MW; A20E08223C537EFE CRC64;

Query Match 77.8%; Score 7; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWFLR 7
Db 97 IETWFLR 103

RESULT 2
MENTO_HUMAN
ID MENTO_HUMAN STANDARD; PRT; 234 AA.
AC Q95772;
DT 15-NOV-2002, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1999, sequence version 1.
DT 07-FEB-2006, entry version 39.
DE MLN64 N-terminal domain homolog (STARD3 N-terminal-like protein).
GN Name=STARD3NL; Synonyms=MENTHO; ORFNames=UNQ855/PRO1864;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, CHARACTERIZATION, TOPOLOGY, PHOSPHORYLATION, AND
RP ALTERNATIVE INITIATION.
RC TISSUE=Petal brain;
RX MEDLINE=22384343; PubMed=12393907; DOI=10.1074/jbc.M208290200;
RA Alpy F., Wendling C., Rio M.-C., Tomasetto C.;
RT "MENTHO, a MLN64 homologue devoid of the START domain."
RL J. Biol. Chem. 277:50780-50787(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heidens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R.L., Watanabe C., Wiand D., Woods K., Xie M.-H.,
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment."
RL Genome Res. 13:2265-2270(2003).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA Fowell G.A., Delahunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Teak A.,
RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,

```



RA Oersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,  
 RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,  
 RA Kozlowski-Reilly A., Leonard S., Rohlfing T., Rock S.M.,  
 RA Tin-Wollam A.M., Abbott A., Minx P., Maupin R., Stowmatt C.,  
 RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,  
 RA Wendt M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,  
 RA Bieri T.A., Nelson J.O., Berkowicz N., Wohlmann P.E., Cook L.B.,  
 RA Hickenbotham M.F., Eldred J., Williams D., Bedell J.A., Mardis E.R.,  
 RA Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,  
 RA Gillett W., Zhou Y., James R., Phelps K., Iadonato S., Bubb K.,  
 RA Stamps E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,  
 RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,  
 RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,  
 RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,  
 RA Waterston R.H., Wilson R.K.;  
 RA "The DNA sequence of human chromosome 7.";  
 RL Nature 424:157-164(2003).  
 RN [4]  
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Brain, and Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Spapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Teshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Myers R.M.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Late endosomal  
 CC membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative initiation;  
 CC Comment=2 isoforms, 1 (shown here) and 2, are produced by  
 CC alternative initiation;  
 CC -1- PTM: Phosphorylated.  
 CC -1- SIMILARITY: Contains 1 MENTAL domain.  
 CC  
 CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC  
 CC -----  
 CC ENBL; A492267; CAD37353.1; -; mRNA.  
 CC ENBL; A358645; AA089008.1; -; mRNA.  
 CC ENBL; AC006033; AA07552.1; -; Genomic\_DNA.  
 CC ENBL; BC003074; AA03074.1; -; mRNA.  
 CC ENBL; BC005959; AA05959.1; -; mRNA.  
 CC Ensembl; ENSG0000010270; Homo sapiens.  
 CC HGNC; HGNC:19169; STARD3NL.  
 KW Alternative initiation; Membrane; Phosphorylation; Transmembrane.  
 FT CHAIN 1 234  
 FT MLN64 N-terminal domain homolog, isoform  
 FT 1.  
 FT CHAIN 8 234  
 FT MLN64 N-terminal domain homolog, isoform  
 FT 2.  
 FT INIT MET 8 8  
 FT TOPO\_DOM 1 53  
 FT Transmem 54 74  
 FT Cytoplasmic (Potential).  
 FT TOPO\_DOM 75 97  
 FT Extracellular (Potential).  
 FT Transmem 98 118  
 FT TOPO\_DOM 119 122  
 FT Cytoplasmic (Potential).

FT TRANSMEM 123 143  
 FT TOPO\_DOM 144 150  
 FT Transmem 151 171  
 FT Transmem 172 234  
 FT Cytoplasmic (Potential).  
 FT DOMAIN 48 218  
 FT MENTAL.  
 SQ SEQUENCE 234 AA; 26655 MW; AFB7DAE381983FB0 CRC64;  
 Query Match 66.7%; Score 6; DB 1; Length 234;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 IETWFL 6  
 Db 166 IETWFL 171  
 RESULT 3  
 MENTO MOUSE  
 ID MENTO MOUSE STANDARD; PRT; 235 AA.  
 AC Q9DCI3; Q99J63; Q9D356;  
 DT 15-NOV-2002, integrated into UniProtKB/Swiss-Prot.  
 DT 15-NOV-2002, sequence version 2.  
 DT 07-FEB-2006, entry version 29.  
 DE MLN64 N-terminal domain homolog (STARD3 N-terminal-like protein).  
 GN Name=Stard3nl; Synonyms=Menthos;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).  
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Kidney;  
 RX PubMed=16141072; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Oyama R., Ravasi F., Lenhard B., Wells C., Kodzius R., Shikawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilming L.G., Ainslie V., Allen J.E.,  
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.I., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Humnietz L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
 RA Petroski Y., Piazza S., Reed J.C., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,  
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Tesdale R.D., Liu E.T., Bruscia V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawashima T., Kojima M., Kondo S., Konno K., Nakano K., Ninomiya N.,  
 RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome.";  
 RL Science 309:1559-1563(2005).



RT Antisense Transcription in the Mammalian Transcriptome.";  
 RL Science 309:1564-1566(2005).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone marrow;  
 RX MEDLINE=22354683; PubMed=12456851; DOI=10.1038/nature01266;  
 RA Okazaki I., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nakaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matuszewska H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Bruscia V., Chochia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer J.S.,  
 RA Gaasterland T., Gariboldi M., Glessi C., Godzik A., Gough J.,  
 RA Gracimond S., Gutinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Portea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Varardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume M., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yashinashi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.,  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone marrow;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Straubli P., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone marrow;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [7]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone marrow;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki P., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RL sequencing pipeline with 384 multicapillary sequencer.";  
 RN Genome Res. 10:1757-1771(2000).  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone marrow;  
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,  
 RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,  
 RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,  
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,  
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,  
 RA Muramatsu M., Hayashizaki Y.,  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; AKL52378; BAE31167.1; -; mRNA.  
 DR MGI; MGI:1923455; Stard3n1.  
 SQ SEQUENCE 235 AA; 26777 MW; AD55148C9FCF73D8 CRC64;  
 Query Match 66.7%; Score 6; DB 2; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IETWFL 6  
 DB 166 IETWFL 171  
 RESULT 5  
 ID Q3U807 MOUSE PRELIMINARY; PRT; 235 AA.  
 AC Q3U807  
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
 DT 11-OCT-2005, sequence version 1.  
 DE B7E-2006, entry version 5.  
 DE Bone marrow macrophage cDNA, RIKEN full-length enriched library,  
 DE clone:IB30047P07 product:STARD3 N-terminal like, full insert sequence.  
 GN Name-Stard3n1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone marrow;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.,  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Methods Enzymol. 303:19-44(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone marrow;  
 RX PubMed=16141072; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilmong L.G., Aizawa S., Allen J.E.,  
 RA Ambesi-Imbriato A., Apweiler R., Aturaliya R.N., Bailey T.L.,  
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA Di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,

```

RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminiecki L., Iacono M., Ieko K., Iwana A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Keiso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottaqui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shinada K., Silva D., Sinclair B.,
RA Sperling S., Stupica E., Sugura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nieuwen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RT Science 309:1559-1563 (2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RC PubMed=16141073; DOI=10.1126/science.1112009;
RX Riken Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RT Science 309:1564-1566 (2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RC MEDLINE=22354693; PubMed=12466851; DOI=10.1038/nature01266;
RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Sato R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Clothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wallestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX

```

---

```

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.F.,
RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbets P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690 (2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RT Genome Res. 10:1617-1630 (2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa K., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RT Genome Res. 10:1757-1771 (2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Inamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DDJB databases.
CC Copyrighted under the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
DR EMBL; AK152116; BAE30960.1; --; mRNA.
DR MGI; MGI:1923455; Stcard3n1.
SQ SEQUENCE 235 AA; 26811 MW; F251725390CB1503 CRC64;
Query Match 66.7%; Score 6; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IETWFL 6
Db 166 IETWFL 171

```

---

```

RESULT 6
Q5U205 RAT PRELIMINARY; PRT; 235 AA.
ID Q5U205_RAT

```

AC Q5U205;  
 DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
 DT 07-DEC-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 10.  
 DE STARD3 N-terminal like (Predicted).  
 DE Name-Stard3n1.  
 GN Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Ovary;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Ovary;  
 RG NIH MGC Project;  
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; BC086352; AAH86352.1; -; mRNA.  
 DR Ensembl; ENSRNOG0000012126; Rattus norvegicus.  
 SQ SEQUENCE 235 AA; 26719 MW; 6238B671397EA775 CRC64;  
  
 Query Match 66.7%; Score 6; DB 2; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 IETWFL 6  
 Db 166 IETWFL 171  
 |||||  
  
 RESULT 7  
 Q4IRY2 GIBZE  
 ID Q4IRY2 GIBZE PRELIMINARY; PRT; 405 AA.  
 AC Q4IRY2;  
 DT 16-AUG-2005, integrated into UniProtKB/TrEMBL.  
 DT 16-AUG-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 4.  
 DE Hypothetical protein.  
 DE OFNames=FG00026.1;  
 GN Gibberella zeae (Fusarium graminearum).  
 OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
 OX NCBI\_TaxID=5518;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=PH-1 / NRRL 31084;

RA Birren B.W., Nusbaum C., Abouelleil A., Allen N., Anderson S.,  
 RA Atachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,  
 RA Boukhalter B., Butler J., Calvo S.E., Camarata J., Chang J.,  
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArelano K.,  
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,  
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D.,  
 RA Galagan J.E., Gardyna S., Gnerre S., Graham L., Grand-Pierre N.,  
 RA Hafez N., Hagopian D., Hagos B., Hall J., Horton L., Hulme W.,  
 RA Iliev I., Jaffe D., Johnson R., Jones C., Kamal M., Kamat A.,  
 RA Karatas A., Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G.,  
 RA Lui A., Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J.,  
 RA Manning J., Matthews C., Maucelli E., McCarthy M., Meldrim J.,  
 RA Meneus L., Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C.,  
 RA Nicol R., Nielsen C.B., Norbu C., O'Connor T., O'Donnell P.,  
 RA O'Neill D., Oliver J., Peterson K., Phunkhang P., Pierre N.,  
 RA Purcell S., Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C.,  
 RA Rogov P., Roman J., Schauer S., Schuback R., Seaman S., Severy P.,  
 RA Smirnov S., Smith C., Spencer B., Stange-Thomann N., Stojanovic N.,  
 RA Stubbs M., Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,  
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,  
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,  
 RA Lander E.S.;  
 RT "Fusarium graminearum genome sequence.";  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; AACM01000002; EAA69365.1; -; Genomic DNA.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 405 AA; 43960 MW; 5E275BB334F6AE3E CRC64;  
  
 Query Match 66.7%; Score 6; DB 2; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 3 TWFLRH 8  
 Db 173 TWFLRH 178  
 |||||  
  
 RESULT 8  
 GP113 HUMAN  
 ID GP113 HUMAN STANDARD; PRT; 1079 AA.  
 AC Q81ZF5; Q6UXT7; Q6UXX3; Q86SL7; Q8IXD8; Q8TDT3;  
 DT 15-FEB-2005, integrated into UniProtKB/Swiss-Prot.  
 DT 01-MAR-2003, sequence version 1.  
 DT 07-MAR-2006, entry version 25.  
 DE Probable G-protein coupled receptor 113 precursor (G-protein coupled  
 DE receptor PGR23).  
 DE Name=GP113; Synonyms=PGR23; ORFNames=UNO9196/PRO34000;  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).  
 RX PubMed=12435584; DOI=10.1016/S0014-5793(02)03574-3;  
 RA Fredriksson R., Lagerstroem M.C., Hoeglund P.J., Schioeth H.B.;  
 RT "Novel human G protein-coupled receptors with long N-terminals  
 RT containing GPS domains and Ser/Thr-rich regions.";  
 RL FEBS Lett. 531:407-414 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
 RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,  
 RA Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;  
 RT "Genome-wide discovery and analysis of human seven transmembrane helix  
 RT receptor genes.";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.





```

RN NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AC010896; AAV14645.1; -; Genomic DNA.
CC Ensembl; ENSG00000173567; Homo sapiens.
CC GO; GO:0016020; C:membrane; IEA.
CC GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
CC GO; GO:0007218; P:neuropeptide signaling pathway; IEA.
CC InterPro; IPR013032; EGF like reg.
CC InterPro; IPR000832; GPCR secretin.
CC InterPro; IPR001879; hormone_rcpt.
CC InterPro; IPR000203; PAD_cys_rich.
CC Pfam; PF00002; 7tm_2; 1.
CC Pfam; PF01825; GPS; 1.
CC PRINTS; PR00249; GPCRSECRETIN.
CC SMART; SM00303; GPS; 1.
CC PROSITE; PS01186; EGF_2; UNKNOWN 1.
CC PROSITE; PS00650; G_PROTEIN_REC_F2_2; UNKNOWN 1.
CC PROSITE; PS02227; G_PROTEIN_REC_F2_3; 1.
CC PROSITE; PS02611; G_PROTEIN_REC_F2_4; 1.
CC PROSITE; PS02221; GPS; 1.
KW Hypothetical protein.
SQ SEQUENCE 1079 AA; 116341 MW; A18CA158F4DDBB9C CRC64;
Query Match 66.7%; Score 6; DB 2; Length 1079;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 WFLRHP 9
DB 267 WFLRHP 272
RESULT 10
Q8XOR0_NEUCR PRELIMINARY; PRT; 1093 AA.
AC Q8XOR0;
DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2002, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Hypothetical protein SE6.080.
GN Name=SE6.080;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Neurospora.
OX NCBI_TaxID=5141;
[1]
RN NUCLEOTIDE SEQUENCE.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RA German Neurospora genome project;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AL670004; CAD21248.1; -; Genomic DNA.

```

```

KW Hypothetical protein.
SQ SEQUENCE 1093 AA; 120695 MW; 9F6BF07A8AD661BD CRC64;
Query Match 66.7%; Score 6; DB 2; Length 1093;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 WFLRHP 9
DB 621 WFLRHP 626
RESULT 11
Q7S795_NEUCR PRELIMINARY; PRT; 1117 AA.
AC Q7S795;
DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT 15-DEC-2003, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Predicted protein.
GN ORFNames=NCU08869.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Neurospora.
OX NCBI_TaxID=5141;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA STRAIN=74-OR23-1A / FGSC 987;
RA MEDLINE=22598136; PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Bell-Pedersen D., Nelson M.A.,
RA Werner-Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L.,
RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,
RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,
RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kanvasselis M.,
RA Mauceli E., Bielke C., Rudd S., Frishman D., Krystofova S.,
RA Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S., Cogoni C.,
RA Macino G., Catchside D.E.A., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.P.C., Glass N.L., Orbach M.J., Berglund J.A., Voelker R.,
RA Yarden O., Flammann M., Seiler S., Dunlap J.C., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.W.;
RL Nature 422:859-868(2003).
CC
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AABX01000300; EAA31426.1; -; Genomic DNA.
SQ SEQUENCE 1117 AA; 123343 MW; 19B8B0ECD607752D CRC64;
Query Match 66.7%; Score 6; DB 2; Length 1117;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 WFLRHP 9
DB 621 WFLRHP 626
RESULT 12
Q6CGE3_YARLI PRELIMINARY; PRT; 29 AA.
AC Q6CGE3;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Similarity.

```

```

GN OrderedLocusNames=YALI02A00429;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CLIB 122 / E 150;
RX PubMed=1529592; DOI=10.1038/nature02579;
RA Dufon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marx C., Neuveglise C., Talla E.,
RA Goffard N., Frangul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykaesten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicoud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in Yeasts.";
RL Nature 430:35-44 (2004).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; CR382127; CAG84207.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 29 AA; 3169 MW; 5FB2DCF7AA4626ED CRC64;

Query Match 55.6%; Score 5; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 15 FLRHP 19

RESULT 13
Q8A2T7_BACTN PRELIMINARY; PRT; 77 AA.
AC Q8A2T7;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Hypothetical protein.
GN OrderedLocusNames=BT3218; ORFNames=BT_3218;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
OX [1];
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=2250858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjuvsell M.K., Hinrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076 (2003).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; A5015928; AAQ78324.1; -; Genomic_DNA.
DR BioCyc; BTHE226186; BT3218-MONOMER; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 77 AA; 9164 MW; 115052AB1896BA18 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 77;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 44 FLRHP 48

RESULT 15
Q46AP0_METBA PRELIMINARY; PRT; 80 AA.
AC Q46AP0;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN OrderedLocusNames=Mbar_A2121;
OS Methanosarcina barkeri.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2208;
OX [1];
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Fusaro / DSM 804;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hammon N., Istrani S., Pittluck S., Goodwin L.A., Saunders E.H.,
RA Schmutz J., Larimer F., Land M., Anderson I., Richardson P.;
RT "Complete sequence of chromosome 1 of Methanosarcina barkeri str.
RT Fusaro.";
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; CP000099; AAZ71052.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 80 AA; 9187 MW; 2DD0ED9A5E5C4CB3 CRC64;

```



Query Match 55.6%; Score 5; DB 2; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 |||||  
 DB 32 FLRHP 36

RESULT 16  
 Q8TIM1\_METAC PRELIMINARY; PRT; 80 AA.  
 AC Q8TIM1\_METAC  
 DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2002, sequence version 1.  
 DT 07-MAR-2006, entry version 11.  
 DE Hypothetical protein.  
 GN ORENAMES=MA\_4126;  
 OS Methanosarcina acetivorans.  
 OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;  
 OC Methanosarcinaceae; Methanosarcina.  
 OX NCBI\_TaxID=2214;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=CZA / ATCC 35395 / DSM 2834;  
 RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;  
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,  
 RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,  
 RA Allen N., Naylor J., Spange-Thomann N., Dearellano K., Johnson R.,  
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,  
 RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,  
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
 RA Springer T.A., Umavay L.A., White O., White R.H., de Macario E.C.,  
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,  
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
 RA Metcalf W.W., Birren B.;  
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
 RT and physiological diversity.";  
 RL Genome Res. 12:532-542(2002).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; AE010299; AAM07474.1; -; Genomic\_DNA.  
 DR GenomeReviews; AE010299 GR; NA4126.  
 DR BioCyc; NA4126-MONOMER; -.  
 DR InterPro; IPR012933; Ycfa.  
 DR Pfam; PF07927; Ycfa; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 80 AA; 9184 MW; DF02721324F5D173 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 |||||  
 DB 32 FLRHP 36

RESULT 17  
 Q5QCQ7\_CENAS PRELIMINARY; PRT; 87 AA.  
 AC Q5QCQ7\_CENAS  
 DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.  
 DT 04-JAN-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 4.  
 DE Hypothetical protein (Fragment).  
 OS Cenibacterium arsenoxidans.  
 OC Bacteria; Cenibacterium.  
 OX NCBI\_TaxID=204773;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ULPasi;  
 RA Carapito C., Muller D., Turlin E., Riegel P., Leize E., Danchin A.,  
 RA Van Dorselaer A., Bertin P., Lett M.-C.; Cenibacterium arsenoxidans, a  
 RT "Pleiotropic effect of arsenic stress on Cenibacterium arsenoxidans, a  
 RT metalloresistant beta-proteobacterium.";  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; AY728027; AAV68356.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 FT NON TER 87  
 SQ SEQUENCE 87 AA; 9653 MW; EAAD40B00A2E3C86 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
 |||||  
 DB 20 WFLRH 24

RESULT 18  
 Q9MIS5\_9TELE PRELIMINARY; PRT; 95 AA.  
 ID Q9MIS5\_9TELE  
 AC Q9MIS5\_9TELE  
 DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.  
 DT 01-OCT-2000, sequence version 1.  
 DT 07-FEB-2006, entry version 21.  
 DE Cytochrome b (Fragment).  
 GN Names=cytb;  
 OS Retropinna tasmanica.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Retropinnidae; Retropinna.  
 OX NCBI\_TaxID=89573;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22111806; PubMed=12116439; DOI=10.1080/106351500750049824;  
 RA Waters J.M., Lopez J.A., Wallis G.P.;  
 RT "Molecular phylogenetics and biogeography of galaxiid fishes  
 RT (Osteichthyes: Galaxiidae): dispersal, vicariance and the position of  
 RT Lepidogalaxias salamandroides.";  
 RL Syst. Biol. 49:777-795(2000).  
 CC -I- FUNCTION: Component of the ubiquinol-cytochrome c reductase  
 CC complex (complex III or cytochrome b-c1 complex), which is a  
 CC coupled to ATP synthesis (By similarity).  
 CC respiratory chain that generates an electrochemical potential  
 CC -I- COFACTOR: Binds 2 heme groups noncovalently (By similarity).  
 CC -I- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,  
 CC cytochrome c1 and the Rieske protein (By similarity).  
 CC -I- SIMILARITY: Belongs to the cytochrome b family.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL; AF112321; AAF67414.1; -; Genomic\_DNA.  
 DR SMR; Q9MIS5; 1-95.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005506; F:iron ion binding; IEA.  
 DR GO; GO:0046872; F:metal ion binding; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR005797; Cytb\_b6\_N.  
 DR Pfam; PF00033; Cytochrom\_B\_N; 1.  
 DR PROSITE; PS51002; CYTB\_NTER; 1.

KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;  
 KW Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 95  
 FT NON\_TER 95 95  
 SQ SEQUENCE 95 AA; 10578 MW; E7F5ABDD28E269DE CRC64;

Query Match 55.6%; Score 5; DB 2; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETWFL 6  
 Db |||||  
 77 ETWFL 81

## RESULT 19

ID Q3BX5\_PELLD PRELIMINARY; PRT; 101 AA.  
 AC Q3BX5;  
 DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.  
 DT 22-NOV-2005, sequence version 1.  
 DE Hypothetical protein.  
 GN ORFNames=Plut 1094;  
 OS Pelodictyon luteolum (strain DSM 273) (Chlorobium luteolum (strain DSM 273)).  
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
 OC Chlorobium/Pelodictyon group; Pelodictyon.  
 OX NCBI\_TaxID=319225;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=DSM 273;  
 RG US DOE Joint Genome Institute;  
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,  
 RA Hammon N., Israni S., Pitluck S., Bryant D., Schmutz J., Larimer F.,  
 RA Land M., Kyrpides N., Ivanova N., Richardson P.,  
 RT "Complete sequence of Pelodictyon luteolum DSM 273."  
 RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; CP000096; ABB23956.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 101 AA; 11173 MW; AAEF2D3DE11B891C CRC64;

Query Match 55.6%; Score 5; DB 2; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
 Db |||||  
 26 TWFLR 30

## RESULT 20

ID Q82LF3\_STRAW PRELIMINARY; PRT; 109 AA.  
 AC Q82LF3;  
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2003, sequence version 1.  
 DE Hypothetical protein.  
 DE Hypothetical protein.  
 GN OrderedLocNames=SAV2057;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN [1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,

RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; BA000030; BAC69768.1; -; Genomic\_DNA.  
 DR BiCyc; SAVE227882:SAV2057-MONOMER; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR InterPro; IPR011991; Wing\_hlx\_DNA\_bd.  
 KW Complete proteome; DNA-binding; Hypothetical protein; Transcription;  
 KW Transcription regulation.  
 SQ SEQUENCE 109 AA; 12127 MW; 9BF1F50C411DFAD2 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 109;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 Db |||||  
 77 FLRHP 81

## RESULT 21

ID Y115\_ADE02 STANDARD; PRT; 115 AA.  
 AC P03290;  
 DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.  
 DT 21-JUL-1986, sequence version 1.  
 DT 07-FEB-2006, entry version 21.  
 DE Hypothetical protein E-115.  
 OS Human adenovirus 2 (HAdV-2).  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=10515;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].  
 RX MEDLINE=83056843; PubMed=7142161;  
 RA Ginges T.R., Sciaky D., Gellinas R.E., Bing-Dong J., Yen C.E.,  
 RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;  
 RT "Nucleotide sequences from the adenovirus-2 genome.";  
 RL J. Biol. Chem. 257:13475-13491(1982).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; J01917; -; NOT\_ANNOTATED\_CDS; Genomic\_DNA.  
 DR PIR; A03862; A03862.  
 KW Hypothetical protein.  
 FT CHAIN 1 115 Hypothetical protein E-115.  
 FT /FTID=PRO\_0000221918.

SQ SEQUENCE 115 AA; 12236 MW; C7A08EA239B8FD98 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETWFL 6  
 Db |||||

```

Db          3  ETWFL 7

RESULT 22
Q2KS22_ADE05
ID  Q2KS22_ADE05  PRELIMINARY;  PRT;  115 AA.
AC  Q2KS22
DT  07-MAR-2006, integrated into UniProtKB/TrEMBL.
DT  07-MAR-2006, sequence version 1.
DT  07-MAR-2006, entry version 1.
DE  Hypothetical 12 kDa early protein.
OS  Human adenovirus 5 (HAdV-5)
OC  Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX  NCBI_TaxID=28285;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=NRRC Ad5FS 7151;
RG  Epidemic Outbreak Surveillance (EOS);
RA  Tibbets C., Purkayastha A., Su J., Russell K., Carlisle S.,
RA  Ospina R., Reynolds T., Rowley R., Hanson E., Seto D.;
RT  "The complete nucleotide sequence and genome organization of Human
RL  adenovirus serotype 5, field strain."
RL  Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
CC  EMBL; AY601635; AAW65500.1; -; Genomic_DNA.
KW  Hypothetical protein.
SQ  SEQUENCE 115 AA; 12210 MW; DF1B2DA39AA7F08 CRC64;

Query Match          55.6%; Score 5; DB 2; Length 115;
Best Local Similarity 100.0%; Pred.No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          2  ETWFL 6
Db          3  ETWFL 7

RESULT 23
Q3WDX0_9ACTO
ID  Q3WDX0_9ACTO  PRELIMINARY;  PRT;  118 AA.
AC  Q3WDX0;
DT  11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT  11-OCT-2005, sequence version 1.
DT  07-FEB-2006, entry version 3.
DE  Putative cytochrome P450.
GN  ORFNames=FraneanIDRAFT_5182;
OS  Frankia sp. EAN1pec.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Frankineae; Frankiaceae; Frankia.
OX  NCBI_TaxID=298653;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=EAN1pec;
RG  US DOE Joint Genome Institute (JGI-PGF);
RA  Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA  Hammon N., Israni S., Pitluck S., Richardson P.;
RT  "Sequencing of the draft genome and assembly of Frankia sp. EAN1pec."
RL  Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RL  [2]
RN  NUCLEOTIDE SEQUENCE.
RC  STRAIN=EAN1pec.
RG  US DOE Joint Genome Institute (JGI-ORNL);
RA  Larimer F., Land M.;
RT  "Annotation of the draft genome assembly of Frankia sp. EAN1pec.";
RL  Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC  -!- CAUTION: The sequence shown here is derived from an
CC  EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC  preliminary data.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms

```

```

CC          3  ETWFL 7
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
CC  EMBL; AAI101000014; EAN16946.1; -; Genomic_DNA.
CC  GO; GO:0020037; F:heme binding; IEA.
CC  GO; GO:0005506; F:iron ion binding; IEA.
CC  GO; GO:0004497; F:monooxygenase activity; IEA.
CC  GO; GO:0006118; P:electron transport; IEA.
CC  InterPro; IPR002397; BP450.
CC  InterPro; IPR001128; Cytochrome_P450.
CC  PRINTS; PR00359; BP450.
SQ  SEQUENCE 118 AA; 13148 MW; 8EEA8775EFE424AD CRC64;

Query Match          55.6%; Score 5; DB 2; Length 118;
Best Local Similarity 100.0%; Pred.No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5  FLRHP 9
Db          61  FLRHP 65

RESULT 24
Q4QKE8_HAB18
ID  Q4QKE8_HAB18  PRELIMINARY;  PRT;  118 AA.
AC  Q4QKE8;
DT  19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT  19-JUL-2005, sequence version 1.
DT  07-FEB-2006, entry version 5.
DE  Putative integrase/recombinase.
GN  OrderedLocusNames=NT11711;
OS  Haemophilus influenzae (strain 86-028NP).
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC  Pasteurellaceae; Haemophilus.
OX  NCBI_TaxID=281310;
RN  [1]
RP  NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX  PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
RA  Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,
RA  Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,
RA  Munson R.S. Jr.;
RT  "Genomic sequence of an otitis media isolate of nontypeable
RT  Haemophilus influenzae: comparative study with H. influenzae serotype
RT  d, strain KW20."
RL  J. Bacteriol. 187:4627-4636(2005).
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
CC  EMBL; CP000057; AAX88499.1; -; Genomic_DNA.
CC  GO; GO:0003677; F:DNA binding; IEA.
CC  GO; GO:0015074; P:DNA integration; IEA.
CC  GO; GO:0006310; P:DNA recombination; IEA.
KW  Complete proteome.
SQ  SEQUENCE 118 AA; 13822 MW; 9FCB660420C82E38 CRC64;

Query Match          55.6%; Score 5; DB 2; Length 118;
Best Local Similarity 100.0%; Pred.No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5  FLRHP 9
Db          21  FLRHP 25

RESULT 25
Q5P5F8_AZOSE
ID  Q5P5F8_AZOSE  PRELIMINARY;  PRT;  122 AA.
AC  Q5P5F8;
DT  04-JAN-2005, integrated into UniProtKB/TrEMBL.
DT  04-JAN-2005, sequence version 1.
DT  07-FEB-2006, entry version 8.
DE  Hypothetical protein.
GN  OrderedLocusNames=AZOSEAL13290; ORFNames=6BA2387;

```

```

OS Azoarcus sp. (strain EbN1).
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Azoarcus.
OX NCBI_TaxID=76114;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15551059; DOI=10.1007/s00203-004-0742-9;
RA Rabus R., Kube M., Heider J., Beck A., Heitmann K., Widdel F.,
RA Reinhardt R.;
RT "The genome sequence of an anaerobic aromatic-degrading denitrifying
bacterium, strain EbN1."
RL Arch. Microbiol. 183:27-36(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CR555306; CA107454.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 122 AA; 13581 MW; 18B790A94ECD3255 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 16 FLRHP 20

RESULT 26
Q47S60_THEFY PRELIMINARY; PRT; 124 AA.
AC Q47S60;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN OrderedLocusNames=Tfu_0689;
OS Thermobifida fusca (strain YX).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptosporangineae; Nocardiopsaceae; Thermobifida.
OX NCBI_TaxID=269800;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG US DOE Joint Genome Institute;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Di Bartolo G., Chain P., Schmutz J.,
RA Larimer F., Land M., Lykidis A., Richardson P.,
RA "Complete sequence of Thermobifida fusca YX."
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000088; AA254707.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 124 AA; 14456 MW; 5F749F9A86A83FC0 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8
DB 115 WFLRH 119

RESULT 27
CRCB_ERWCT STANDARD; PRT; 127 AA.
AC Q6D7N0;
DT 05-JUL-2005, integrated into UniProtKB/Swiss-Prot.
DT 16-AUG-2004, sequence version 1.

```

```

DT 07-MAR-2006, entry version 15.
DE Protein crcB homolog.
GN Name=crcB; OrderedLocusNames=ECA1295;
OS Erwinia carotovora subsp. atroseptica (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebatina M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
carotovora subsp. atroseptica and characterization of virulence
factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- SUBCELLULAR LOCATION: Bacterial cell inner membrane; multi-pass
membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the crcB family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BX950851; CAG74205.1; -; Genomic DNA.
DR GenomeReviews; BX950851_GR; ECA1295.
DR HAMAP; MF_00454; -; 1.
DR InterPro; IPR003691; Camphor_CrcB.
DR Pfam; PF02537; CRCB; 1.
DR TIGRFAMs; TIGR00494; crcB; 1.
KW Complete proteome; inner membrane; Membrane; Transmembrane.
FT CHAIN 1 127
FT PROTEIN CRcB homolog.
FT TRANSMEM 4 24
FT TRANSMEM 35 55
FT TRANSMEM 71 91
FT TRANSMEM 103 123
FT SEQUENCE 127 AA; 13391 MW; ADED63C701397633 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 55 FLRHP 59

RESULT 28
Q90Z26_XENTR PRELIMINARY; PRT; 128 AA.
ID Q90Z26;
AC Q90Z26;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Xcat-2.
GN Name=Xcat-2;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Vempati U.D., King M.L.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License

```

```

CC EMBL; AF256086; AAK49295.1; -; mRNA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006445; P:regulation of translation; IEA.
DR InterPro; IPR008705; Nanos_RNA_bd.
DR Pfam; PF05741; zf-nanos; I.
SQ SEQUENCE 128 AA; 14140 MW; E79556DEF1C080B8B CRC64;

Query Match 55.6%; Score 5; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9
Db 117 FLRHP 121

RESULT 29
Q2SYI4_BURTH
ID Q2SYI4_BURTH PRELIMINARY; PRT; 130 AA.
AC Q2SYI4;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Hypothetical protein.
GN ORFNames=BTH_I1318;
OS Burkholderia thailandensis E264.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=271848;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=E264;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Palmer N., Dodson R.,
RA Hickey E.K., Gwinn M., Dougherty B., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van-Vugt R., Adams M.D., Gocayne J.D., Weidman J., Uterback T.,
RA Watthey L., McDonald L., Artiach P., Bowman C., Garland S., Fujii C.,
RA Cotton M.D., Horst K., Tomb J.-F., Roberts K., Hatch B., Smith H.O.,
RA Venter J.C.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; CP000086; ABC36802.1; -; Genomic_DNA.
DR Hypothetical protein.
SQ SEQUENCE 130 AA; 14536 MW; D2998BDEE2181906 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWF 5
Db 12 IETWF 16

RESULT 30
Q4HDP5_CAMCO
ID Q4HDP5_CAMCO PRELIMINARY; PRT; 130 AA.
AC Q4HDP5;
DT 16-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Hypothetical protein.
GN ORFNames=CCOA0060;
OS Campylobacter coli RM2228.
OX Plasmid pCC178.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.

```

```

OX NCBI_TaxID=306254;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2228;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AF01000016; EAL55992.1; -; Genomic_DNA.
DR Hypothetical protein; Plasmid.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 130 AA; 15007 MW; DDF9E96C0D0262FC CRC64;

Query Match 55.6%; Score 5; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWF 5
Db 5 IETWF 9

RESULT 31
Q6TIW1_ANETH
ID Q6TIW1_ANETH PRELIMINARY; PRT; 132 AA.
AC Q6TIW1;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 16.
DE Putative transposase.
OS Aneurinibacillus thermoaerophilus.
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae;
OC Aneurinibacillus group; Aneurinibacillus.
OX NCBI_TaxID=143495;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L420-917;
RX PubMed=15044388; DOI=10.1093/glycob/cwh064;
RA Schaffer C., Messner P.;
RT "Surface-layer glycoproteins: an example for the diversity of
RT bacterial glycosylation with promising impacts on nanobiotechnology.";
RL Glycobiology 14:31R-42R(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L420-917;
RX PubMed=15316277;
RA Novotny R., Pfeostl A., Messner P., Schaffer C.;
RT "Genetic organization of chromosomal S-layer glycan biosynthesis loci
RT of Bacillaceae.";
RL Glycoconj. J. 20:435-447(2004).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AY442352; AAS55727.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR InterPro; IPR012337; RNaseH_fold.
KW Hydrolase; Nuclease.
SQ SEQUENCE 132 AA; 15708 MW; 603062293C9D57B0 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;

```

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 TWELR 7
Db 25 TWELR 29
RESULT 32
Q62H5 BURMA
ID Q62H5_BURMA PRELIMINARY; PRT; 133 AA.
AC Q62H5;
DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein.
GN OrderedLocusNames=BMA2316;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.B., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR TIGR; BMA2316; -; Genomic_DNA.
DR InterPro; IPR011944; CHP2246.
DR TIGRFAMs; TIGR02246; Cons hypoth_2246; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 133 AA; 14849 MW; F61094D42700C13E CRC64;
Query Match 55.6%; Score 5; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IETWF 5
Db 12 IETWF 16
RESULT 33
Q63R57 BURPS
ID Q63R57_BURPS PRELIMINARY; PRT; 133 AA.
AC Q63R57;
DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Hypothetical protein.
GN OrderedLocusNames=BPSL2816;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Tibball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA
```

```
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crossett B., Davis P., DeShazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; BX571965; CAH36826.1; -; Genomic_DNA.
DR InterPro; IPR011944; CHP2246.
DR TIGRFAMs; TIGR02246; Cons hypoth_2246; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 133 AA; 14819 MW; F6065863EAA5913E CRC64;
Query Match 55.6%; Score 5; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IETWF 5
Db 12 IETWF 16
RESULT 34
Q411L0 KINRA
ID Q411L0_KINRA PRELIMINARY; PRT; 134 AA.
AC Q411L0;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=KradDRAFT_2276;
OS Kineococcus radiotolerans SRS30216.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Frankineae; Kineosporiaceae; Kineococcus.
OX NCBI_TaxID=266940;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SRS30216;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hamon N., Israni S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Kineococcus
RT radiotolerans SRS30216.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SRS30216;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of draft genome assembly of Kineococcus radiotolerans
RT SRS30216.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SRS30216;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hamon N., Israni S., Pittluck S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC
CC !- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
```

```
DR EMBL; AABF02000024; EAM74977.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 134 AA; 15058 MW; F037BBE97A0D4676 CRC64;

Query Match
Best Local Similarity 55.6%; Score 5; DB 2; Length 134;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TWFLR 7
Db 41 TWFLR 45

RESULT 35
Q977K8_9CREN
ID Q977K8_9CREN PRELIMINARY; PRT; 136 AA.
AC Q977K8;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Secreted protein.
OS uncultured crenarchaeote 74A4.
OC Archaea; Crenarchaeota; environmental samples;
OC marine archaeal group 1.
OX NCBI_TaxID=166279;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21633832; PubMed=11772643; DOI=10.1128/ASM.68.1.335-345.2002;
RA Beja O., Koonin E.V., Aravind L., Taylor L.T., Seitz H., Stein J.L.,
RA Bensen D.C., Feldman R.A., Swanson R.V., DeLong E.F.;
RT "Comparative Genomic Analysis of Archaeal Genotypic Variants in a
RL Single Population and in Two Different Oceanic Provinces.";
RA Appl. Environ. Microbiol. 68:335-345(2002).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AFJ93466; AAK96100.1; -; Genomic_DNA.
SQ SEQUENCE 136 AA; 15922 MW; 852D6DD1B1626B5C CRC64;

Query Match
Best Local Similarity 55.6%; Score 5; DB 2; Length 136;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TWFLR 7
Db 61 TWFLR 65

RESULT 36
Q82HV8_STRAW
ID Q82HV8_STRAW PRELIMINARY; PRT; 136 AA.
AC Q82HV8;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Hypothetical protein.
GN OrderedLocusNames=SAV3400;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites.";
RA Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

[2]
RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RA Nat. Biotechnol. 21:526-531(2003).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; BA000030; BAC71112.1; -; Genomic_DNA.
DR BioCyc; SAV227882:SAV3400-MONOMER; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 136 AA; 14797 MW; 8A1E1A1D59C1F6F3 CRC64;

Query Match
Best Local Similarity 55.6%; Score 5; DB 2; Length 136;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9
Db 11 FLRHP 15

RESULT 37
O83574_TREPA
ID O83574_TREPA PRELIMINARY; PRT; 137 AA.
AC O83574;
DT 01-NOV-1998, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1998, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Hypothetical protein.
GN OrderedLocusNames=TP0563; ORFNames=TP_0563;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artiach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Mitochondrial; inner
CC membrane (By similarity).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AE000520; AAC65546.1; -; Genomic_DNA.
DR PIR; A71308; A71308.
DR TIGR; TP0563; -.
DR GO; GO:0031072; F:heat shock protein binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0008457; P:protein folding; IEA.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ; 1.
DR SMART; SM00271; DnaJ; 1.
KW Chaperone; Complete proteome; Hypothetical protein; Inner membrane;
KW Membrane; Protein transport; Translocation; Transmembrane; Transport.
SQ SEQUENCE 137 AA; 15724 MW; AF812BC67B76F2E CRC64;

Query Match
Best Local Similarity 55.6%; Score 5; DB 2; Length 137;
```



```

Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWF 5
DB 122 IETWF 126

RESULT 38
ID Q73P38 TREDE PRELIMINARY; PRT; 137 AA.
AC Q73P38;
DT 05-JUL-2004, integrated into UniProtKB/TREMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE DnaJ domain protein.
GN OrderedLocusNames=TDE0961; ORFNames=TDE_0961;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Sehadri R., Myers G.S.A., Tetelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidson T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
RA Gebregeorgis E., Geer K., Tsagaye G., Malek J.A., Ayodeji B.,
RA Shatman S., McLeod M.P., Snajs D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstein G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT with other spirochete genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Mitochondrial; inner
CC membrane (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
EMBL: AF017226; AAS11452.1; -; Genomic_DNA.
DR TIGR; TDE0961; -;
DR BioCyc; TDE243275:TDE0961-MONOMER; -;
DR GO; GO:0031072; F:heat shock protein binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ; 1.
DR SMART; SM00271; DnaJ; 1.
KW Chaperone; Complete proteome; Inner membrane; Membrane;
KW Protein transport; Translocation; Transmembrane; Transport.
SQ SEQUENCE 137 AA; 15752 MW; D2AA37F990E7D4EC CRC64;

Query Match 55.6%; Score 5; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IETWF 5
DB 120 IETWF 124

RESULT 39
ID Q61GY3 DROME PRELIMINARY; PRT; 139 AA.
AC Q61GY3;
DT 05-JUL-2004, integrated into UniProtKB/TREMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE HDC04272.
GN ORFNames=HDC04272;
OS Drosophila melanogaster (Fruit fly).

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovyev V., Busold C.,
RA Feilenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J.D.,
RA Faro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
EMBL: BK003633; DAA02331.1; -; Genomic DNA.
DR EMBL; BK003633; DAA02331.1; -; Genomic DNA.
SQ SEQUENCE 139 AA; 14909 MW; DBA4F99D58E4045D CRC64;

Query Match 55.6%; Score 5; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 78 FLRHP 82

RESULT 40
QANBA0_9M1CC
ID Q4NBA0_9M1CC PRELIMINARY; PRT; 139 AA.
AC Q4NBA0;
DT 19-JUL-2005, integrated into UniProtKB/TREMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-MAR-2006, entry version 5.
DE Similar to Glutaredoxin and related proteins.
GN ORFNames=ArthDRAPT_0289;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
EMBL: AAHG0100023; EAL94631.1; -; Genomic DNA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR011515; Glix_actino.
DR InterPro; IPR012336; Thioroxn-like_fd.
DR InterPro; IPR006662; Thiorox.
DR InterPro; IPR006663; Thioroxox_dom2.

```



DR PRINTS; PR00421; THIOREDOXIN.  
DR TIGRFAMS; TIGR02200; G1rX actino; 1.  
DR PROSITE; PS00194; THIOREDOXIN; UNKNOWN 1.  
SQ SEQUENCE 139 AA; 14842 MW; 188B6C4668B8D6E6 CRC64;  
Query Match 55.6%; Score 5; DB 2; Length 139;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FLRHP 9  
| | | | |  
Db 19 FLRHP 23  
Search completed: August 31, 2006, 10:39:41  
Job time : 140.25 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:29:54 ; Search time 107.75 Seconds  
(without alignments)  
38.190 Million cell updates/sec

Title: DENGUE\_SEROTYPE3

Perfect score: 9

Sequence: 1 retwflrhp 9

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2589679 seqs, 457216429 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2564502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : A Geneseq\_8:\*

- 1: Geneseqp1990s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*
- 9: Geneseqp2005s:\*
- 10: Geneseqp2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	88.9	39	9	Adw12582 M1-40/DEN
2	8	88.9	48	9	Adw12588 P(95-114)
3	6	66.7	278	8	Adq25888 Human GPC
4	6	66.7	826	5	Abb07253 Human nov
5	6	66.7	827	6	Abu07568 Human sec
6	6	66.7	904	4	Abg09947 Novel hum
7	6	66.7	924	5	Aab71323 Human GCR
8	6	66.7	953	7	Ade34415 Human G-p
9	6	66.7	994	5	Abb07252 Human nov
10	6	66.7	994	5	Aau99808 Novel hum
11	6	66.7	994	7	Ade34425 Human G-p
12	6	66.7	994	8	Ado28977 Human nov
13	6	66.7	994	8	Adq25892 Human gua
14	6	66.7	1018	5	Aae25061 Human G-p
15	6	66.7	1070	6	Abu07567 Human sec
16	6	66.7	1131	4	Abg11655 Novel hum
17	6	66.7	1232	7	Adf70474 Orphan re
18	5	55.6	28	10	Aee37134 Human ser
19	5	55.6	34	4	Aau17769 Novel hum
20	5	55.6	34	7	Adga41149 Human res
21	5	55.6	34	7	Adi96923 Human res
22	5	55.6	52	6	Abg99963 Human nov
23	5	55.6	60	4	Aau42843 Propionib

24	5	55.6	60	6	ABM39362	Abm39362 Propionib
25	5	55.6	62	4	AAm86926	AAm86926 Human imm
26	5	55.6	64	4	AAg98737	AAg98737 Human cel
27	5	55.6	64	4	AAU50032	AAU50032 Propionib
28	5	55.6	64	6	ABM46551	ABM46551 Propionib
29	5	55.6	73	5	ABP03674	ABP03674 Human ORF
30	5	55.6	75	4	AAg98736	AAg98736 Human cel
31	5	55.6	76	4	AAm99844	AAm99844 Human exc
32	5	55.6	76	4	AAm42659	AAm42659 Human kid
33	5	55.6	82	5	ABP34862	ABP34862 Human ORF
34	5	55.6	85	8	ADT58131	ADT58131 Plant pol
35	5	55.6	86	4	AAU51518	AAU51518 Propionib
36	5	55.6	86	6	ABM48037	ABM48037 Propionib
37	5	55.6	89	8	ADX94950	ADX94950 Plant ful
38	5	55.6	98	9	ABM94143	ABM94143 M. xanthu
39	5	55.6	102	7	ADC14235	ADC14235 Human enz
40	5	55.6	104	2	AAU74113	AAU74113 Human pro
41	5	55.6	106	4	AAU09103	AAU09103 Novel hum
42	5	55.6	108	4	AAO07214	AAO07214 Human pol
43	5	55.6	110	5	ABU51319	ABU51319 Helicobac
44	5	55.6	111	6	ABP75900	ABP75900 Human sec
45	5	55.6	124	8	ADX90486	ADX90486 Plant ful
46	5	55.6	125	9	AEA79622	AEA79622 IC6 Mab h
47	5	55.6	125	9	AEC39351	AEC39351 Human IC6
48	5	55.6	127	4	AAm50235	AAm50235 Catalpa s
49	5	55.6	127	5	AAU76417	AAU76417 Catalpa l
50	5	55.6	146	5	ADK36828	ADK36828 Novel hum
51	5	55.6	149	8	ADY23809	ADY23809 Plant ful
52	5	55.6	154	2	AAW07588	AAW07588 Fibroblas
53	5	55.6	154	2	AAW07587	AAW07587 Fibroblas
54	5	55.6	154	2	AAW07589	AAW07589 Fibroblas
55	5	55.6	154	2	AAW07590	AAW07590 Fibroblas
56	5	55.6	154	3	AAV90462	AAV90462 Mutant hu
57	5	55.6	154	3	AAV90464	AAV90464 Mutant hu
58	5	55.6	154	3	AAV90460	AAV90460 Saporin e
59	5	55.6	154	3	AAV90461	AAV90461 Mutant hu
60	5	55.6	154	3	AAV90463	AAV90463 Mutant hu
61	5	55.6	154	7	ADC34717	ADC34717 Human fib
62	5	55.6	154	7	ADC34715	ADC34715 Human fib
63	5	55.6	154	7	ADC34718	ADC34718 Human fib
64	5	55.6	154	7	ADC34716	ADC34716 Human fib
65	5	55.6	154	7	ADH92147	ADH92147 Fibroblas
66	5	55.6	154	7	ADH92148	ADH92148 Fibroblas
67	5	55.6	154	7	ADH92145	ADH92145 Codon opt
68	5	55.6	154	7	ADH92146	ADH92146 Fibroblas
69	5	55.6	158	5	ABU51724	ABU51724 Helicobac
70	5	55.6	175	4	AAU64446	AAU64446 Propionib
71	5	55.6	175	6	ABM60965	ABM60965 Propionib
72	5	55.6	179	4	AAE06639	AAE06639 Human alp
73	5	55.6	188	9	ABM92796	ABM92796 M. xanthu
74	5	55.6	189	4	AAm93280	AAm93280 Human pol
75	5	55.6	189	8	ADL30724	ADL30724 Human pro
76	5	55.6	198	6	ABR58403	ABR58403 Human nov
77	5	55.6	201	5	ABR90287	ABR90287 Human nov
78	5	55.6	215	9	AED00443	AED00443 Lactobaci
79	5	55.6	216	4	AAm39930	AAm39930 Human pol
80	5	55.6	225	9	AED00317	AED00317 Lactobaci
81	5	55.6	234	4	AAU29056	AAU29056 Human PRO
82	5	55.6	234	4	AAU39929	AAU39929 Human pol
83	5	55.6	234	4	ABR87532	ABR87532 Human PRO
84	5	55.6	234	5	ABG95857	ABG95857 Human sec
85	5	55.6	234	5	ABR84847	ABR84847 Human PRO
86	5	55.6	234	5	ABR95453	ABR95453 Human PRO
87	5	55.6	234	6	ABU58432	ABU58432 Human PRO
88	5	55.6	234	6	ABU87980	ABU87980 Novel hum
89	5	55.6	234	6	ABU84295	ABU84295 Human sec
90	5	55.6	234	6	ABR66169	ABR66169 Human sec
91	5	55.6	234	6	ABR65559	ABR65559 Human sec
92	5	55.6	234	6	ABU99499	ABU99499 Human sec
93	5	55.6	234	6	ABU82738	ABU82738 Human PRO
94	5	55.6	234	6	ABU89859	ABU89859 Novel hum
95	5	55.6	234	6	ABR68108	ABR68108 Human sec
96	5	55.6	234	6	ABU96161	ABU96161 Novel hum

```
97 5 55.6 234 6 ABU92592 Human sec
98 5 55.6 234 6 ABO08669 Human sec
99 5 55.6 234 6 ABO02721 Human sec
100 5 55.6 234 6 ABR74875 Human sec
101 5 55.6 234 6 ABR94637 Human sec
102 5 55.6 234 6 ABU85610 Human PRO
103 5 55.6 234 6 ABU98770 Human PRO
104 5 55.6 234 6 ABU97985 Novel hum
105 5 55.6 234 6 ABU91691 Novel hum
106 5 55.6 234 6 ABU89384 Human PRO
107 5 55.6 234 6 ABU86225 Human sec
108 5 55.6 234 6 ABU67438 Human sec
109 5 55.6 234 6 ABU80466 Human PRO
110 5 55.6 234 6 ABU90882 Novel hum
111 5 55.6 234 6 ABO033941 Human sec
112 5 55.6 234 6 ABR99384 Human sec
113 5 55.6 234 6 ABR98774 Human sec
114 5 55.6 234 6 ABO16297 Human sec
115 5 55.6 234 6 ABR92197 Human sec
116 5 55.6 234 6 ABO18838 Human sec
117 5 55.6 234 6 ABR78259 Human sec
118 5 55.6 234 6 ABR39937 Human PRO
119 5 55.6 234 6 ABU71958 Novel hum
120 5 55.6 234 6 ABU84995 Novel hum
121 5 55.6 234 6 ABO00134 Novel hum
122 5 55.6 234 6 ABO11466 Human sec
123 5 55.6 234 6 ABO02111 Human sec
124 5 55.6 234 6 ABU88685 Novel hum
125 5 55.6 234 6 ABU83380 Human sec
126 5 55.6 234 6 ABO06181 Novel hum
127 5 55.6 234 6 ABR59217 Human sec
128 5 55.6 234 6 ABO09279 Human sec
129 5 55.6 234 6 ABO19143 Novel hum
130 5 55.6 234 6 ABO11161 Human sec
131 5 55.6 234 6 ABR66779 Human sec
132 5 55.6 234 6 ABO15992 Human sec
133 5 55.6 234 6 ABO13698 Human sec
134 5 55.6 234 6 ABU71512 Human sec
135 5 55.6 234 6 ABU65601 Human sec
136 5 55.6 234 6 ABO07449 Human PRO
137 5 55.6 234 6 ABO03636 Human sec
138 5 55.6 234 6 ABR67084 Human sec
139 5 55.6 234 6 ABO15687 Human sec
140 5 55.6 234 6 ABU55968 Human sec
141 5 55.6 234 6 ABU72293 Human PRO
142 5 55.6 234 6 ABU65296 Human PRO
143 5 55.6 234 6 ABU95241 Novel hum
144 5 55.6 234 6 ABU71144 Human PRO
145 5 55.6 234 6 ABO07754 Human PRO
146 5 55.6 234 6 ABR69995 Human sec
147 5 55.6 234 6 ABR69328 Human sec
148 5 55.6 234 6 ABO01469 Human PRO
149 5 55.6 234 6 ABU81271 Human PRO
150 5 55.6 234 6 ABR60068 Human sec
```

## ALIGNMENTS

```
RESULT 1
ADW12582
ID ADW12582 standard; peptide; 39 AA.
XX
AC ADW12582;
XX
DT 24-MAR-2005 (first entry)
XX
XX M1-40/DEN-2 (F36) mutant protein.
XX
DE Gene therapy; protein purification; virucide; cytostatic; vaccine;
KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX DEN; dengue; mutant; mutein.
XX
```

```
OS Dengue virus.
XX US2004266987-A1.
FN 30-DEC-2004.
XX
XX 30-JUN-2003; 2003US-00608029.
PF 30-JUN-2003; 2003US-00608029.
XX
PR 30-JUN-2003; 2003US-00608029.
XX (INSP ) INST PASTEUR.
XX
PI Despres P, Catteau A;
XX WPI; 2005-047647/05.
XX
XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
PT as a vaccine for preventing or treating pathological conditions from non-
PT specific febrile illnesses to severe hemorrhagic manifestations or
PT encephalitic syndromes.
XX
XX Example 1; SEQ ID NO 29; 30pp; English.
XX
XX The present invention relates to an isolated and purified ApoptoM
CC peptide. The invention is useful as a vaccine for the prevention and
CC treatment of pathological conditions from non-specific febrile illnesses
CC to severe hemorrhagic manifestations, encephalitic syndromes and these
CC pathological conditions are linked to flavivirus infection or cancers.
CC The invention is also useful in gene therapy. The present sequence is a
CC M1-40/DEN (dengue)-2 (F36) mutant protein.
XX
XX Sequence 39 AA;
SQ
Query Match 88.9%; Score 8; DB 9; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ETWFLRHP 9
Db 32 ETWFLRHP 39
RESULT 2
ADW12588
ID ADW12588 standard; protein; 48 AA.
XX
XX ADW12588;
AC 24-MAR-2005 (first entry)
XX
XX p(95-114) EGFP(M1-M40)DEN-2 (136F) plasmid DNA encoded protein #3.
XX
XX Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
KW DEN; dengue; EGFP; enhanced green fluorescent protein.
XX
XX Dengue virus.
OS Chimeric.
OS Unidentified.
XX
XX Key Location/Qualifiers
FH Misc-difference 2 /note= "Encoded by GGC"
FT Misc-difference 4 /note= "Encoded by GAC"
FT Misc-difference 13.44 /note= "Encoded by GTTTC"
XX
XX US2004266987-A1.
XX
XX 30-DEC-2004.
XX
XX 30-JUN-2003; 2003US-00608029.
PF
```

```

XX PR 30-JUN-2003; 2003US-00608029.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX XX WPI; 2005-047647/05.
XX DR N-PSDB; ADW12589.
XX XX
XX PT New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
XX PT as a vaccine for preventing or treating pathological conditions from non-
XX PT specific febrile illnesses to severe hemorrhagic manifestations or
XX PT encephalitic syndromes.
XX XX
XX PS Disclosure; SEQ ID NO 35; 30pp; English.
XX XX
XX CC The present invention relates to an isolated and purified ApoptoM
XX CC peptide. The invention is useful as a vaccine for the prevention and
XX CC treatment of pathological conditions from non-specific febrile illnesses
XX CC to severe hemorrhagic manifestations, encephalitic syndromes and these
XX CC pathological conditions are linked to flavivirus infection or cancers.
XX CC The invention is also useful in gene therapy. The present sequence is a
XX CC p(95-114) EGFP (enhanced green fluorescent protein) (M1-W40)DEN (Dengue)-2
XX CC (136F) plasmid DNA encoded protein.
XX XX
XX SQ Sequence 48 AA;
XX
XX Query Match 88.9%; Score 8; DB 9; Length 48;
XX Best Local Similarity 100.0%; Pred. No. 0.036;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 ETWFLRHP 9
XX DB 41 ETWFLRHP 48
XX
XX RESULT 3
XX ADQ25888
XX ID ADQ25888 standard; protein; 278 AA.
XX AC ADQ25888;
XX XX
XX DT 23-SEP-2004 (first entry)
XX XX
XX DE Human GPCR related protein #1.
XX XX
XX KW receptor; GPCR; guanosine triphosphate-binding protein-coupled receptor;
XX KW human.
XX OS Homo sapiens.
XX XX
XX PN WO2004055186-A1.
XX XX
XX PD 01-JUL-2004.
XX PF 18-DEC-2003; 2003WO-JP016245.
XX XX
XX PR 18-DEC-2002; 2002JP-00366417.
XX PR 03-MAR-2003; 2003JP-00055691.
XX XX
XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX XX
XX PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX XX
XX DR WPI; 2004-500216/47.
XX DR N-PSDB; ADQ25887.
XX XX
XX PT New polynucleotide encoding guanosine triphosphate-binding protein-
XX PT coupled receptor, for use in developing a therapeutic agent for medical
XX PT treatment.
XX XX

```

---

```

PS XX Example 5; SEQ ID NO 16; 104pp; Japanese.
XX
XX CC The present invention provides the protein and coding sequences of a
XX CC human guanosine triphosphate-binding protein-coupled receptor (GPCR). The
XX CC sequences are useful for treating diseases related to the abnormality of
XX CC the expression of GPCR, and for developing a therapeutic agent for
XX CC medical treatment. The present sequence is a protein shown in the
XX CC exemplification of the invention.
XX
XX SQ Sequence 278 AA;
XX
XX Query Match 66.7%; Score 6; DB 8; Length 278;
XX Best Local Similarity 100.0%; Pred. No. 35;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 4 WFLRHP 9
XX DB 185 WFLRHP 190
XX
XX RESULT 4
XX ABB07253
XX ID ABB07253 standard; protein; 826 AA.
XX XX
XX AC ABB07253;
XX XX
XX DT 26-MAR-2002 (first entry)
XX XX
XX DE Human novel GPCR (NGPCR) protein.
XX XX
XX KW G coupled protein receptor; GPCR; NGPCR; cytostatic; anorectic; cancer;
XX KW antiinflammatory; immunosuppressive; antidiabetic; human.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200187932-A2.
XX XX
XX PD 22-NOV-2001.
XX XX
XX PF 11-MAY-2001; 2001WO-US015048.
XX XX
XX PR 12-MAY-2000; 2000US-0203875P.
XX PR 30-MAY-2000; 2000US-0207932P.
XX XX
XX PA (LEXI-) LEXICON GENETICS INC.
XX XX
XX PI Hu Y, Nepomnichy B, Wang X, Walke DW, Gerhardt B, Turner CA;
XX XX
XX DR WPI; 2002-114231/15.
XX DR N-PSDB; ABA94352.
XX XX
XX PT New polypeptide, useful for generation of antibodies and for screening
XX PT compounds for treatment of mental, biological or medical disorders and
XX PT diseases, comprises the isolated G coupled protein receptor polypeptide.
XX XX
XX PS Claim 8; Page 81-83; 85pp; English.
XX XX
XX CC The invention provides novel G coupled protein receptor (GPCR) proteins
XX CC and polynucleotides encoding the same. The novel GPCR (NGPCR) proteins
XX CC can be expressed by standard recombinant methodology. The NGPCR proteins
XX CC and polynucleotides are useful for diagnosis, in treatment of diseases,
XX CC drug screening, clinical trial monitoring, for treatment of physiological
XX CC or behavioural disorders, for the detection of mutant GPCRs or
XX CC inappropriately expressed GPCR for the diagnosis of disease, and for
XX CC screening drugs effective in the treatment of the symptomatic or
XX CC phenotypic manifestations of perturbing the normal function of GPCR in
XX CC the body. The NGPCR proteins are useful for the generation of antibodies,
XX CC as reagents in diagnostic assays, for the identification of other
XX CC cellular gene products related to a GPCR, as reagents in assays for
XX CC screening compounds that can be used as pharmaceutical reagents for the
XX CC therapeutic treatment of mental, biological or medical disorders and
XX CC diseases, and for identifying compounds useful in the therapeutic
XX CC treatment of obesity, inflammation, immune disorders, diabetes, heart and

```

CC coronary disease, metabolic disorders, and cancer. The present sequence  
 CC represents a human NGPCR protein  
 XX  
 SQ Sequence 826 AA;  
 Query Match 66.7%; Score 6; DB 5; Length 826;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRHP 9  
 Db 21 WFLRHP 26  
 RESULT 5  
 ABU07568  
 ID ABU07568 standard; protein; 827 AA.  
 XX  
 AC ABU07568;  
 XX  
 DT 20-MAR-2003 (first entry)  
 XX  
 DE Human secretin type G protein-coupled receptor #2.  
 KW Human; receptor; GPCR; G protein-coupled receptor; secretin; obesity;  
 KW cardiovascular disorder; diabetes; infection; HIV; pain; cancer;  
 KW human immunodeficiency virus infection; anorexia; bulimia; asthma;  
 KW Parkinson's disease; acute heart failure; hypotension; hypertension;  
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;  
 KW ulcer; allergy; benign prostatic hypertrophy; psychosis;  
 KW neurological disorder; anxiety; schizophrenia; manic depression;  
 KW delirium; dementia; mental retardation; dyskinesia; Huntington's disease;  
 KW Tourette's syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 827  
 FT /note= "Encoded by GA"  
 XX  
 PN WO200299106-A2.  
 XX  
 PD 12-DEC-2002.  
 XX  
 PF 03-JUN-2002; 2002WO-EP006031.  
 XX  
 PR 04-JUN-2001; 2001US-0294998P.  
 PR 26-JUL-2001; 2001US-0307608P.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Koehler RH, Smolyar A;  
 XX  
 DR WPI; 2003-140623/13.  
 DR N-PSDB; ABX15279.  
 XX  
 PT New isolated polynucleotide encoding human secretin-type G protein-  
 PT coupled receptor (GPCR) polypeptides, useful for preventing or treating  
 PT diseases associated with GPCR dysfunction, e.g. cardiovascular disease or  
 PT diabetes.  
 XX  
 PS Claim 1; Fig 7; 127pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide which: (a) encodes a  
 CC human secretin-type G protein-coupled receptor (GPCR) polypeptide; (b)  
 CC comprises a sequence appearing as ABX15278 and ABX15279; (c) hybridises  
 CC under stringent conditions to the polynucleotide in (A) and (B); (d) has  
 CC a sequence deviating from (A)-(C) due to the degeneration of the genetic  
 CC code, or represents a fragment, derivative or allelic variation of (A)-  
 CC (D). Also included are an expression vector containing the above  
 CC polynucleotide, a host cell containing the expression vector, a  
 CC substantially purified human secretin-type GPCR polypeptide, methods of  
 CC screening for agents which modulate or decrease the activity of a human

CC secretin-type GPCR, methods of reducing the activity of the human  
 CC secretin-type GPCR, the identified modulators. The polynucleotide is  
 CC useful in preventing, ameliorating, or treating diseases associated with  
 CC human secretin-type GPCR dysfunction. The polynucleotide may also be used  
 CC as hybridisation probes or primers, and in diagnostic assays or in  
 CC genetic testing. The methods are useful in producing and detecting the  
 CC polynucleotide and polypeptide and in screening for agents that modulate  
 CC the activity of the human secretin-type GPCR. The expression vector or  
 CC the reagent is useful in preparing a medicament for modulating the  
 CC activity of a human secretin-type GPCR in a disease, such as a  
 CC cardiovascular disorder, obesity, diabetes, infections (bacterial, viral,  
 CC fungal and protozoan), HIV (human immunodeficiency virus) infection,  
 CC pain, cancer, anorexia, bulimia, asthma, Parkinson's disease, acute heart  
 CC failure, hypotension, hypertension, urinary retention, osteoporosis,  
 CC angina pectoris, myocardial infarction, ulcers, allergies, benign  
 CC prostatic hypertrophy, psychosis, neurological disorders (e.g. anxiety,  
 CC schizophrenia, manic depression, delirium, dementia, mental retardation,  
 CC dyskinesias, Huntington's disease and Tourette's syndrome). The present  
 CC sequence represents a human secretin type GPCR of the invention  
 XX  
 SQ Sequence 827 AA;  
 Query Match 66.7%; Score 6; DB 6; Length 827;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRHP 9  
 Db 21 WFLRHP 26  
 RESULT 6  
 ABG09947  
 ID ABG09947 standard; protein; 904 AA.  
 XX  
 AC ABG09947;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #9938.  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2..  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS74134.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 40306; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 904 AA;  
 SQ

Query Match 66.7%; Score 6; DB 4; Length 904;  
 Best Local Similarity 100.0%; Pred. No. 96;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRHP 9  
 Db 156 WFLRHP 161  
 |||||

RESULT 7  
 AAB71323  
 ID AAB71323 standard; protein; 924 AA.  
 XX  
 AC AAB71323;  
 XX  
 DT 19-NOV-2002 (first entry)  
 XX  
 DE Human GCRC-2 INCYTE ID 7474890CD1 SEQ ID 2.  
 XX  
 KW GCRC; Human; G-protein coupled receptor; anti-HIV; antiarteriosclerotic;  
 KW cyostatic; neuroprotective; antiparkinsonian; hepatotropic; laxative;  
 KW cerebroprotective; antiinflammatory; virucide; antibacterial; fungicide;  
 KW protozoacide; cirrhosis; cancer; stroke; Alzheimer's disease; AIDS;  
 KW Parkinson's disease; Crohn's disease; constipation; infection; receptor;  
 KW gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200263004-A2.  
 PN  
 XX 15-AUG-2002.  
 PD  
 XX 06-FEB-2002; 2002WO-US003635.  
 PF  
 XX 07-FEB-2001; 2001US-0267322P.  
 PR 23-FEB-2001; 2001US-0271215P.  
 PR 08-MAR-2001; 2001US-0274551P.  
 PR 23-MAR-2001; 2001US-0278507P.  
 PR 30-MAR-2001; 2001US-0280597P.  
 PR 02-APR-2001; 2001US-0281107P.  
 PR 06-APR-2001; 2001US-0282121P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 XX Baughn MR, Tribouley CM, Nguyen DB, Thornton M, Yao MG;  
 PI Kallick DA, Gandhi AR, Wallia NK, Arvizu C, Elliott VS, Hafalia AJA;  
 PI Ramkumar J, Pei J, Tang YT, Yue H, Reddy R, Butford N, Lu DAM;  
 PI Graul RC, Khan FA, Walsh RT, Ison CH, Richardson TW, Griffin JA;  
 PI Warren BA, Yang J, Lee EA, Harland L;  
 XX  
 DR WPI; 2002-627557/67.  
 DR N-PSDB; AAF88581.  
 XX

PT New human G-protein coupled receptors (GCRC), useful for diagnosing or  
 PT treating a disease or condition associated with decreased expression or  
 PT over expression of functional GCRCs e.g. cancer, Alzheimer's and  
 PT Parkinson's.  
 XX  
 PS Claim 63; Page 160-163; 239pp; English.  
 XX  
 CC This invention describes novel polypeptides which have anti-HIV,  
 CC antiarteriosclerotic, cyostatic, neuroprotective, antiparkinsonian,  
 CC hepatotropic, laxative, cerebroprotective, antiinflammatory, virucide,  
 CC antibacterial, fungicide and protozoacide activity. The products of the  
 CC invention are useful for treating a disease or condition associated with  
 CC decreased expression or over expression of functional G-protein coupled  
 CC receptors (GCRC), while antibodies generated against the polypeptide of  
 CC the invention are useful for diagnosing a condition or disease associated  
 CC with the expression of GCRC e.g. arteriosclerosis, cirrhosis, cancer,  
 CC stroke, Alzheimer's disease, Parkinson's disease, Crohn's disease,  
 CC constipation, AIDS, or bacterial, viral, fungal or protozoal infections.  
 CC The compounds described in the invention can be used for gene therapy.  
 CC AAB71322-AAB71369 represent the GCRC proteins encoded by AAF88580-  
 CC AAF88627 described in the disclosure of the invention  
 XX  
 SQ Sequence 924 AA;  
 Query Match 66.7%; Score 6; DB 5; Length 924;  
 Best Local Similarity 100.0%; Pred. No. 98;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRHP 9  
 Db 141 WFLRHP 146  
 |||||

RESULT 8  
 ADE34415  
 ID ADE34415 standard; protein; 953 AA.  
 XX  
 AC ADE34415;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human G-protein coupled receptor protein #SEQ ID 35.  
 XX  
 KW Cytostatic; antiinflammatory; hepatotropic; nephrotropic; dermatological;  
 KW antiarthritic; antiasthmatic; antidiabetic; hypotensive; antiulcer;  
 KW antilipemic; antiarteriosclerotic; neurotropic; neuroprotective; anorectic;  
 KW immunomodulator; uropathic; antiinfertility; G-protein coupled receptor;  
 KW GPCR; GPCR185; GPCR186; GPCR187; GPCR188; GPCR189; GPCR222; GPCR223;  
 KW hepatitis; nephritis; dermatitis; pancreatitis; rheumatoid arthritis;  
 KW osteoarthritis; atopic dermatitis; asthma; diabetes; hypertension;  
 KW inflammatory bowel disease; gastric ulcer; arteriosclerosis;  
 KW hyperlipemia; Alzheimer's disease; dementia; obesity; pulmonary fibrosis;  
 KW renal fibrosis; immune deficiency; infertility; urinary blockage; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003078632-A1.  
 PN  
 XX 25-SEP-2003.  
 PD  
 XX 14-MAR-2003; 2003WO-JP003050.  
 PF  
 XX 15-MAR-2002; 2002JP-00071567.  
 PR 14-MAY-2002; 2002JP-00138013.  
 PR 28-FEB-2003; 2003JP-00054663.  
 XX  
 XX (NISB ) JAPAN TOBACCO INC.  
 PA  
 XX Watanabe H, Nozaki Y;  
 PI  
 XX WPI; 2003-722435/68.  
 DR  
 XX G-protein coupled receptor proteins, genes encoding them and antibodies

PT recognizing them for treatment and diagnosis of cancer, inflammatory and  
 PT gastrointestinal disorders.

PS Example: SEQ ID NO 35; 274pp; Japanese.

XX The invention relates to G-protein coupled receptor proteins of human  
 CC origin. These proteins include GPCR185, GPCR186, GPCR187, GPCR188,  
 CC GPCR189, GPCR222 and GPCR223. Proteins of the invention are used in the  
 CC treatment and prevention of diseases associated with inflammation,  
 CC angiogenesis and tissue neogenesis, including hepatitis, nephritis,  
 CC dermatitis, pancreatitis, rheumatoid arthritis, osteoarthritis, atopic  
 CC dermatitis, asthma, diabetes, hypertension, inflammatory bowel disease,  
 CC gastric ulcer, arteriosclerosis, hyperlipemia, Alzheimer's disease,  
 CC dementia, obesity, pulmonary fibrosis, renal fibrosis, immune deficiency,  
 CC infertility, urinary blockage and cancer (such as cancer of the brain,  
 CC neck, tongue, lung, breast, pancreas, stomach, colon, duodenum, prostate,  
 CC bladder, ovary, womb or rectum). Primers of the invention are devised and  
 CC synthesised based on G-protein coupled receptor consensus sequences and  
 CC used for 5'-RACE (rapid amplification of cDNA ends) and 3'-RACE  
 CC amplification of human cDNA derived from adrenal and visual cortex RNA.  
 CC Sequences given in ADE34534-ADE34533 represent human G-protein coupled  
 CC receptor proteins, genes encoding them, and primers for the amplification  
 CC of these sequences.

XX SQ Sequence 953 AA;

Query Match 66.7%; Score 6; DB 7; Length 953;  
 Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0;

QY 4 WFLRHP 9  
 DB 141 WFLRHP 146  
 |||||

RESULT 9

ABB07252  
 ID ABB07252 standard; protein; 994 AA.

AC ABB07252;

XX 26-MAR-2002 (first entry)

DE Human novel GPCR (NGPCR) protein.

XX G coupled protein receptor; GPCR; NGPCR; cytostatic; anorectic; cancer;  
 KW antiinflammatory; immunosuppressive; antidiabetic; human.

XX OS Homo sapiens.

XX PN WO200187932-A2.

XX PD 22-NOV-2001.

XX PF 11-MAY-2001; 2001WO-US015048.

XX PR 12-MAY-2000; 2000US-0203875P.

XX PR 30-MAY-2000; 2000US-0207932P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Hu Y, Nepomnichy B, Wang X, Walke DW, Gerhardt B, Turner CA;

XX DR WPI; 2002-114231/15.

XX DR N-PSDB; ABA94351.

XX New polypeptide, useful for generation of antibodies and for screening  
 PT compounds for treatment of mental, biological or medical disorders and  
 PT diseases, comprises the isolated G coupled protein receptor polypeptide.

XX Claim 8; Page 78-80; 85pp; English.

XX The invention provides novel G coupled protein receptor (GPCR) proteins

CC and polynucleotides encoding the same. The novel GPCR (NGPCR) proteins  
 CC can be expressed by standard recombinant methodology. The NGPCR proteins  
 CC and polynucleotides are useful for diagnosis, in treatment of diseases,  
 CC drug screening, clinical trial monitoring, for treatment of physiological  
 CC or behavioural disorders, for the detection of mutant GPCRs or  
 CC inappropriately expressed GPCR for the diagnosis of disease, and for  
 CC screening drugs effective in the treatment of the symptomatic or  
 CC phenotypic manifestations of perturbing the normal function of GPCR in  
 CC the body. The NGPCR proteins are useful for the generation of antibodies,  
 CC as reagents in diagnostic assays, for the identification of other  
 CC cellular gene products related to a GPCR, as reagents in assays for  
 CC screening compounds that can be used as pharmaceutical reagents for the  
 CC therapeutic treatment of mental, biological or medical disorders and  
 CC diseases, and for identifying compounds useful in the therapeutic  
 CC treatment of obesity, inflammation, immune disorders, diabetes, heart and  
 CC coronary disease, metabolic disorders, and cancer. The present sequence  
 CC represents a human NGPCR protein

SQ Sequence 994 AA;

Query Match 66.7%; Score 6; DB 5; Length 994;  
 Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Gaps 0;  
 Matches 6; Conservative 0;

QY 4 WFLRHP 9  
 DB 189 WFLRHP 194  
 |||||

RESULT 10

AAU99808

ID AAU99808 standard; protein; 994 AA.

AC AAU99808;

XX 07-OCT-2002 (first entry)

XX Novel human G protein-coupled receptor hTGR21-1.

XX Human; G protein-coupled; receptor; hTGR21; central nervous disease;  
 KW endocrine disease; metabolic disease; cancer; inflammation; nontropic;  
 KW circulatory disorder; respiratory disorder; digestive disorder;  
 KW immune system disorder; infection; gene therapy; neuroprotective;  
 KW antiinflammatory; immunomodulator; cardiant; antimicrobial; cytostatic;  
 XX gene therapy; hTGR21-1.

XX OS Homo sapiens.

XX PN WO200253593-A1.

XX PD 11-JUL-2002.

XX PF 27-DEC-2001; 2001WO-JP011530.

XX PR 28-DEC-2000; 2000JP-00400625.

XX PR 13-APR-2001; 2001JP-00115916.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Miwa M, Ito T, Shintani Y, Miyajima N;

XX DR WPI; 2002-528854/56.

XX DR N-PSDB; ABK88069.

XX Human kidney-originated G protein-coupled receptor protein hTGR21 and  
 PT encoding DNA, for developing drugs to treat e.g. central nervous  
 PT diseases, endocrine diseases, inflammations and diseases of digestive  
 PT system.

XX Claim 1; Page 106-110; 143pp; Japanese.

XX The invention describes a novel human kidney-originated G protein-coupled  
 CC receptor protein hTGR21 and the DNA encoding it. The proteins, DNAs and





The invention relates to human and mouse G protein-coupled receptors (GPCRs) and nucleic acids encoding them. The invention also relates to sequences at least 90% identical to the GPCR proteins and nucleic acids of the invention; methods of treating, preventing or diagnosing diseases associated with GPCRs of the invention; methods of screening for compounds useful in the treatment of GPCR-related diseases; a transgenic mouse comprising a GPCR gene of the invention; a mouse comprising a mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived from the transgenic mice; kits comprising several mice, each of which has a mutation in a different GPCR gene of the invention; and kits comprising probes which hybridise to GPCR polynucleotides of the invention. The invention further discloses variants of the GPCR polypeptides and vectors comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may be used in the diagnosis, treatment or prevention of a wide variety of diseases including neurological disorders (e.g., Alzheimer's disease, depression, diabetic neuropathy, Parkinson's disease or schizophrenia); disorders of the adrenal gland; disorders of the colon or intestine (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or myocardial infarction); muscular disorders; blood disorders (e.g., anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g., obesity, enzyme deficiency-related diseases or vitamin deficiency-related diseases); and disorders of the kidney, liver, lung, breast, ovary, uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and thyroid (e.g., cancers). The present sequence represents a GPCR of the invention. Note: The full sequence data for this patent did not form part of the printed specification; those sequences not shown were obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 994 AA;

Query Match 66.7%; Score 6; DB 8; Length 994;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9

DB 189 WFLRHP 194

RESULT 13

ADQ25892

ID ADQ25892 standard; protein; 994 AA.

XX AC ADQ25892;

XX DT 23-SEP-2004 (first entry)

XX DE Human guanosine triphosphate-binding protein-coupled receptor.

XX KW receptor; GPCR; guanosine triphosphate-binding protein-coupled receptor; human.

XX KW Homo sapiens.

XX OS WO2004055186-A1.

XX PN 01-JUL-2004.

XX PD 18-DEC-2003; 2003WO-JP016245.

XX PF 18-DEC-2002; 2002JP-00366417.

XX PR 03-MAR-2003; 2003JP-00055691.

XX PA (NAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

XX PI Suwa M, Asai K, Akiyama Y, Aburatani H;

XX WPI; 2004-500216/47.

DR N-PSDB; ADQ25891.

XX New polynucleotide encoding guanosine triphosphate-binding protein-coupled receptor, for use in developing a therapeutic agent for medical treatment.

PS Claim 1; SEQ ID NO 20; 104pp; Japanese.

XX The present invention provides the protein and coding sequences of a human guanosine triphosphate-binding protein-coupled receptor (GPCR). The sequences are useful for treating diseases related to the abnormality of the expression of GPCR, and for developing a therapeutic agent for medical treatment. The present sequence is the protein of the invention.

XX Sequence 994 AA;

Query Match 66.7%; Score 6; DB 8; Length 994;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9

DB 189 WFLRHP 194

RESULT 14

AAE25061

ID AAE25061 standard; protein; 1018 AA.

XX AC AAE25061;

XX DT 30-OCT-2002 (first entry)

XX DE Human G-protein coupled receptor (GCRC)-1 protein.

XX KW Human; G-protein coupled receptor; GCRC; olfactory; taste sensation; cell proliferative disorder; actinic keratosis; leukaemia; metabolic; epilepsy; Alzheimer's disease; cardiovascular; hypertension; virucide; angina pectoris; myocardial infarction; gastrointestinal; anorexia; cholecystitis; Crohn's disease; inflammatory; hypotensive; cardiac; acquired immune deficiency syndrome; anaemia; asthma; hepatocytic; diabetes; obesity; infection; transgenic; gene therapy; cytostatic; anticonvulsant; neuroprotective; antiinflammatory; neurological; nootropic; anorectic; autoimmune; receptor.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..27

FT Peptide /label= Signal\_peptide

FT Peptide 1..16

FT Peptide /label= Signal\_peptide

FT Protein 17..1018

FT Protein /note= "Human mature GCRC-1 protein"

FT Protein 28..1018

FT Protein /note= "Human mature GCRC-1 protein"

FT Domain 657..710

FT Domain /note= "Latrophilin/CL-1-like GPS domain"

XX WO200246230-A2.

XX PN 13-JUN-2002.

XX PD 05-DEC-2001; 2001WO-US046659.

XX PF 08-DEC-2000; 2000US-0254323P.

XX PR 13-DEC-2000; 2000US-0255564P.

XX PR 21-DEC-2000; 2000US-025716P.

XX PR 19-JAN-2001; 2001US-0262848P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Kallick DA, Baughn MR, Lu DAM, Yue H, Graul RC, Lu Y, Ding L;

```

PI Tribouley CM, Tang YT, Gandhi AR, Thornton M;
XX WPI; 2002-519657/55.
XX N-PSDB; AAD40625.
XX Novel isolated human G-protein coupled receptor protein useful for
XX diagnosing, treating, preventing hypertension, myocardial infarction,
XX anorexia, cholecystitis, anemia, asthma, diabetes, obesity, Alzheimer's
XX disease.
XX Claim 1; Page 117-120; 136pp; English.
XX The invention relates to human G-protein coupled receptors (GPEC) and
XX their corresponding nucleic acids. GPEC is useful in screening for
XX compounds which acts as its agonist or antagonist and is also useful for
XX preparing a polyclonal or monoclonal antibody. GPEC is useful for
XX identifying a compound that modulates, mimics and/or blocks an olfactory
XX and/or taste sensation. GPEC DNA is useful for assessing toxicity of a
XX test compound. GPEC and its DNA are useful in the diagnosis, treatment
XX and prevention of a cell proliferative disorder e.g. actinic keratosis,
XX leukaemia, etc., a neurological disorder e.g. epilepsy, Alzheimer's
XX disease, etc., a cardiovascular disorder e.g. hypertension, angina
XX pectoris, myocardial infarction, etc., a gastrointestinal disorder e.g.
XX anorexia, cholecystitis, Crohn's disease, etc., an autoimmune/
XX inflammatory disorder e.g. acquired immune deficiency syndrome, anaemia,
XX asthma, etc., a metabolic disorder e.g. diabetes, obesity, etc., and an
XX infection by a viral agent such as adenovirus, arenavirus, etc. GPEC DNA
XX is used for creating knock out or knock in humanised animals or
XX transgenic animals to model human diseases, and somatic or germline gene
XX therapy for treating the above mentioned disorders. The present sequence
XX is human GPEC-1 protein
SQ Sequence 1018 AA;

Query Match 66.7%; Score 6; DB 5; Length 1018;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9
Db 213 WFLRHP 218
|||||

RESULT 15
ABU07567
ID ABU07567 standard; protein; 1070 AA.
XX AC ABU07567;
XX DT 20-MAR-2003 (first entry)
XX DE Human secretin type G protein-coupled receptor #1.
XX KW Human; receptor; GPCR; G protein-coupled receptor; secretin; obesity;
XX cardiovascular disorder; diabetes; infection; HIV; pain; cancer;
XX human immunodeficiency virus infection; anorexia; bulimia; asthma;
XX Parkinson's disease; acute heart failure; hypotension; hypertension;
XX urinary retention; osteoporosis; angina pectoris; myocardial infarction;
XX ulcer; allergy; benign prostatic hypertrophy; psychosis;
XX neurological disorder; anxiety; schizophrenia; manic depression;
XX delirium; dementia; mental retardation; dyskinesia; Huntington's disease;
XX Tourette's syndrome.
XX OS Homo sapiens.
XX PN WO200299106-A2.
XX PD 12-DEC-2002.
XX PF 03-JUN-2002; 2002WO-EP006031.
XX PR 04-JUN-2001; 2001US-0294998P.
XX PR 26-JUL-2001; 2001US-0307608P.

(FARB ) BAYER AG.
Koehler RH, Smolyar A;
WPI; 2003-140623/13.
N-PSDB; ABX15278.
New isolated polynucleotide encoding human secretin-type G protein-
coupled receptor (GPCR) polypeptides, useful for preventing or treating
diseases associated with GPCR dysfunction, e.g. cardiovascular disease or
diabetes.
Claim 1; Fig 2; 127pp; English.
The invention relates to an isolated polynucleotide which: (a) encodes a
human secretin-type G protein-coupled receptor (GPCR) polypeptide; (b)
comprises a sequence appearing as ABX15278 and ABX15279; (c) hybridises
under stringent conditions to the polynucleotide in (A) and (B); (d) has
a sequence deviating from (A)-(C) due to the degeneration of the genetic
code; or represents a fragment, derivative or allelic variation of (A)-
(D). Also included are an expression vector containing the above
polynucleotide, a host cell containing the expression vector, a
substantially purified human secretin-type GPCR polypeptide, methods of
screening for agents which modulate or decrease the activity of a human
secretin-type GPCR, methods of reducing the activity of the human
secretin-type GPCR, the identified modulators. The polynucleotide is
useful in preventing, ameliorating, or treating diseases associated with
human secretin-type GPCR dysfunction. The polynucleotide may also be used
as hybridisation probes or primers, and in diagnostic assays or in
genetic testing. The methods are useful in producing and detecting the
polynucleotide and polypeptide and in screening for agents that modulate
the activity of the human secretin-type GPCR. The expression vector or
the reagent is useful in preparing a medicament for modulating the
activity of a human secretin-type GPCR in a disease, such as a
cardiovascular disorder, obesity, diabetes, infections (bacterial, viral,
fungal and protozoan), HIV (human immunodeficiency virus) infection,
pain, cancer, anorexia, bulimia, asthma, Parkinson's disease, acute heart
failure, hypotension, hypertension, urinary retention, osteoporosis,
angina pectoris, myocardial infarction, ulcers, allergies, benign
prostatic hypertrophy, psychosis, neurological disorders (e.g. anxiety,
schizophrenia, manic depression, delirium, dementia, mental retardation,
dyskinesias, Huntington's disease and Tourette's syndrome). The present
sequence represents a human secretin type GPCR of the invention
XX SQ Sequence 1070 AA;

Query Match 66.7%; Score 6; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9
Db 258 WFLRHP 263
|||||

RESULT 16
ABG11655
ID ABG11655 standard; protein; 1131 AA.
XX AC ABG11655;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #11646.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PR

```

PD 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 XX 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 XX DR N-PSDB; AAS75842.  
 DR New isolated polynucleotide and encoded polypeptides, useful in  
 XX diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PT Claim 20; SEQ ID NO 42014; 103pp; English.  
 PS This invention relates to isolated polynucleotide (I) and polypeptide (II)  
 XX sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABO0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 1131 AA;  
 SQ Query Match 66.7%; Score 6; DB 4; Length 1131;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 4 WFLRHP 9  
 DB 213 WFLRHP 218  
 RESULT 17  
 ADF70474  
 ID ADF70474 standard; protein; 1232 AA.  
 XX ADF70474;  
 AC ADF70474;  
 XX 12-FEB-2004 (first entry)  
 DT Orphan receptor ligand-related human protein SeqID97.  
 DE ligand; orphan receptor protein; fusion protein; fluorescent protein;  
 XX cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
 KW GFPuv; Enhanced GFP; EGFP; human.  
 KW Homo sapiens.  
 OS WO2003071272-A1.  
 XX 28-AUG-2003.  
 PD

XX 21-FEB-2003; 2003WO-JP001901.  
 XX 22-FEB-2002; 2002JP-00045728.  
 PR 23-JUL-2002; 2002JP-00213949.  
 PR 11-OCT-2002; 2002JP-00298237.  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;  
 XX WPI; 2003-697654/66.  
 PI N-PSDB; ADF70576.  
 DR Transformation of cells with a fusion protein of an orphan receptor  
 PT protein with a fluorescent protein useful for identification of ligands  
 PT to the orphan receptor.  
 PT Disclosure; SEQ ID NO 97; 594pp; Japanese.  
 PS This invention relates to a novel method of identifying ligands to an  
 XX orphan receptor protein which comprises transforming cells with DNA  
 CC encoding a fusion protein of the orphan receptor with a fluorescent  
 CC protein, so that the fusion protein is expressed in the cells (or cell  
 CC membranes isolated from them) and contacting the cells with the potential  
 CC ligand to be tested. A suitable fluorescent protein for incorporation in  
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
 CC identification of ligands binding to an orphan receptor protein.  
 XX Sequence 1232 AA;  
 SQ Query Match 66.7%; Score 6; DB 7; Length 1232;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 4 WFLRHP 9  
 DB 189 WFLRHP 194  
 RESULT 18  
 ABE37134  
 ID ABE37134 standard; peptide; 28 AA.  
 XX ABE37134;  
 AC ABE37134;  
 XX 09-FEB-2006 (first entry)  
 DT Human serum N-linked glycopeptide SEQ ID NO: 1238.  
 DE Bioinformatics; blood; serum; plasma protein; protein detection;  
 KW mass spectroscopy; proteomics; glycosylation; diagnosis; cancer;  
 KW cystostatic; diabetes; antidiabetic; inflammation; antiinflammatory;  
 KW rheumatoid arthritis; antiarthritic; antirheumatic; psychiatric disorder;  
 KW neuroleptic; neurological disease; infection; antimicrobial.  
 XX Homo sapiens.  
 OS WO2005114221-A2.  
 XX 01-DEC-2005.  
 PD 20-MAY-2005; 2005WO-US017842.  
 PF 21-MAY-2004; 2004US-0573593P.  
 PR (SYST-) INST SYSTEMS BIOLOGY.  
 XX Aebersold RH, Zhang H;  
 PI WPI; 2006-020173/02.  
 XX

PT Identifying glycopolypeptides in a serum or plasma sample, by identifying  
PT released sample glycopeptide fragments that correspond to standard  
peptides.

PS Claim 1; SEQ ID NO 1238; 193pp; English.

XX The invention relates to identifying glycopolypeptides in a serum or  
CC plasma sample comprising immobilizing derivatized sample  
CC glycopolypeptides to a solid support, releasing the sample glycopeptide  
CC fragments from the solid support, adding to the released sample  
CC glycopeptide fragments standard peptides, and identifying released sample  
CC glycopeptide fragments that correspond to standard peptides added by mass  
CC spectroscopy. Also included are a method for identifying one or more  
CC diagnostic markers for a disease, a composition comprising peptides  
CC containing the glycosylation sites (ABE35897-AEE39378, where the peptides  
CC each correspond to peptide fragments derived by cleavage of polypeptides  
CC using the same cleavage reagent) and a kit comprising peptides containing  
CC the glycosylation sites (ABE35897-AEE39378). The methods are useful for  
CC identifying glycopolypeptides in a serum or plasma sample. The methods  
CC can be used for blood serum profiling for the detection of prognostic and  
CC diagnostic protein markers. It can also be used to identify and/or  
CC validate drug targets and to evaluate drug efficacy, drug dosing, and/or  
CC drug toxicity. The methods can also be used for the detection of changes  
CC in the state of glycosylation of proteins based on the concurrent  
CC application of protein abundance measurement of protein glycosylation on  
CC the same sample. The method allows fast throughput and simplicity. It can  
CC be readily adapted for high throughput analysis of samples, which can be  
CC particularly advantageous for the analysis of clinical specimens. The  
CC method can also be automated to facilitate the processing of multiple  
CC samples. The present sequence is a human glycopeptide comprising an N-  
CC linked glycosylation site, suitable for use as a reference peptide in the  
CC method of the invention.

XX Sequence 28 AA;

Query Match 55.6%; Score 5; DB 10; Length 28;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
|  
|  
|  
|  
Db 14 FLRHP 18

RESULT 19

AAU17769  
ID AAU17769 standard; protein; 34 AA.

XX AAU17769;

AC AAU17769;

XX 07-NOV-2001 (first entry)

XX Novel human respiratory antigen #85.

XX Human, respiratory antigen; respiratory disorder; throat disorder;  
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;  
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;  
KW respiratory active.

XX Homo sapiens.

XX WO20015448-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001333.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-021486P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0234484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251903P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-476224/51.  
DR N-PSDB; AAS27953.  
XX  
XX Isolated polypeptide for treating, preventing and/ or prognosing  
PT disorders related to the respiratory system including respiratory cancers  
PT and also for testing and detection e.g. diagnosis.  
XX  
XX Claim 11; SED ID No 387; 546pp; English.  
PS  
XX The present invention relates to the isolation of novel human respiratory  
CC antigens, and cDNA (AAS27869-AAS28159) and genomic sequences encoding for

CC these polypeptides. The sequences of the invention are useful for  
CC preventing, treating and/or prognosing disorders related to the  
CC respiratory system including throat disorders (e.g. vocal cord paralysis,  
CC tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic  
CC disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose  
CC disorders and cancers of the respiratory tissues e.g. lung cancer. The  
CC polynucleotide sequences of the invention are useful in gene therapy and  
CC anisense therapy. AAU17685-AAU17975 represent novel human respiratory  
CC antigens. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 34 AA;  
  
Query Match 55.6%; Score 5; DB 4; Length 34;  
Best Local Similarity 100.0%; Pred.No. 85;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4 WFLRH 8  
Db 28 WFLRH 32  
|||||  
  
RESULT 20  
ADG41149  
ID ADG41149 standard; protein; 34 AA.  
XX AC ADG41149;  
XX DT 26-FEB-2004 (first entry)  
XX DE Human respiratory system associated protein seq id 387.  
XX  
XX antiinflammatory; antiallergic; antiasthmatic; cytostatic; gene therapy;  
KW respiratory system antigen;  
KW human respiratory system associated polynucleotide;  
KW respiratory system disorder; throat disorder; vocal cord paralysis;  
KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;  
KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;  
KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;  
KW cancer; respiratory tissue cancer; throat cancer; lung cancer;  
KW cancer of the nose; gene therapy; chromosome identification; forensic;  
KW human respiratory system associated protein; human.  
XX  
XX Homo sapiens.  
XX US2003215893-A1.  
XX  
XX 20-NOV-2003.  
XX  
XX 07-AUG-2002; 2002US-00212872.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.

PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226686P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	03-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	03-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231124P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-02311243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231141P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	14-SEP-2000;	2000US-02331968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-02332397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-02332398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-02332399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234274P.	PR	08-DEC-2000;	2000US-0251990P.
PR	25-SEP-2000;	2000US-0234597P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234598P.	PR	05-JAN-2001;	2001US-0259678P.
PR	26-SEP-2000;	2000US-0235484P.	PR	17-JAN-2001;	2001US-00764860.
PR	27-SEP-2000;	2000US-0235834P.	PR	14-FEB-2002;	2002US-00074095.
PR	27-SEP-2000;	2000US-0235836P.	XX	(HUMA-) HUMAN GENOME SCI INC.	
PR	28-SEP-2000;	2000US-0235935P.	XX	Rosen CA, Ruben SM, Barash SC;	
PR	28-SEP-2000;	2000US-0236327P.	PI	WPI; 2003-902033/82.	
PR	29-SEP-2000;	2000US-0236367P.	XX	N-PSDB; ADG40857.	
PR	29-SEP-2000;	2000US-0236368P.	DR	Novel respiratory system antigen and polynucleotides encoding the	
PR	29-SEP-2000;	2000US-0236369P.	DR		

CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders  
 CC (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic  
 CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,  
 CC sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or  
 CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and  
 CC cancer of the nose). The polynucleotides are useful in gene therapy

Query Match 55.6%; Score 5; DB 7; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 WFLRH 8  
 Db 28 WFLRH 32

RESULT 21  
 ADI96923  
 ID ADI96923 standard; peptide; 34 AA.

XX AC ADI96923;

XX DT 04-NOV-2004 (first entry)

XX DE Human respiratory system associated polypeptide SeqID387.

XX KW respiratory system-related polypeptide; antiasthmatic; antibacterial;  
 KW antiinflammatory; cycostatic; antianaemic; antiallergic; gene therapy;  
 KW pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;  
 KW anaemia; leukaemia; inflammation; sinusitis;  
 KW chronic obstructive pulmonary disease; infectious disease; human.

OS Homo sapiens.

XX PN US2003077704-A1.

XX PD 24-APR-2003.

XX PF 14-FEB-2002; 2002US-00074095.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0188874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0232403P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235835P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239335P.  
 PR 13-OCT-2000; 2000US-0239337P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.



PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 11-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 PR 17-JAN-2001; 2001US-00764860.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

WPI: 2003-765403/72.

N-PSDB; ADI96631.

New human respiratory system-related polypeptide and genes, useful for treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or sinusitis.

Claim 11; SEQ ID NO 387; 202pp; English.

This invention is related to a novel isolated polypeptide, which comprises a human respiratory system-related polypeptide, and the DNA sequence which encodes it. The invention may be useful for the development of compounds with an antiasthmatic, antibacterial, anti-inflammatory, cytostatic, antianaemic or antiallergic activity. In addition, the sequences disclosed may be useful for gene therapy. The polypeptide or polynucleotide is useful for treating, preventing or ameliorating a medical condition, for example pneumonia, lung cancer, cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia, inflammations, sinusitis, chronic obstructive pulmonary disease or infectious diseases. The polypeptide or polynucleotide is also useful for diagnosing any of these diseases or a susceptibility to the disease. The present sequence is that of a human respiratory system associated polypeptide of the invention.

Sequence 34 AA;

Query Match 55.6%; Score 5; DB 7; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRH 8  
 Db 28 WFLRH 32

RESULT 22

ABG99963  
 ID ABG99963 standard; protein; 52 AA.

AC ABG99963;

DT 17-JAN-2003 (first entry)

DE Human novel polypeptide #76.

KW Human; genetic disorder; gene mapping; medical imaging; cancer; neurodegenerative disorder; lymphoid cell disorder; osteoporosis; Parkinson's disease; Alzheimer's disease; bone degenerative disorder; osteoarthritis; periodontal disease; liver fibrosis; viral infection; fungal infection; bacterial infection; autoimmune disease; diabetes; atopic dermatitis.

OS Homo sapiens.

PN WO200274961-A1.

PD 26-SEP-2002.

PF 14-MAR-2002; 2002WO-US005109.

PR 15-MAR-2001; 2001US-00810173.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;

DR WPI: 2003-040556/03.

DR N-PSDB; ABX05061.

PT New isolated polypeptides and polynucleotides, useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders, lymphoid cell disorders, bone degenerative disorders, and infections.

PS Claim 9; SEQ ID NO 602; 235pp; English.

CC The invention relates to human polynucleotides and the polypeptides they encode. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, medical imaging, identification of mutations, responsible for genetic disorders or other traits, assessing biodiversity and producing many other types of data and products dependent on DNA and amino acid sequences. They are also useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell disorders, osteoporosis, osteoarthritis, bone degenerative disorders, periodontal disease, liver fibrosis, infections (e.g. viral, fungal or bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis). Sequences ABG9988-ABG9989 and ABU0010-ABU00433 represent human polypeptides of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

Sequence 52 AA;

Query Match 55.6%; Score 5; DB 6; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
 Db 33 FLRHP 37

RESULT 23  
 AAU42843  
 ID AAU42843 standard; protein; 60 AA.  
 XX

AC AAU42843;  
 XX 27-FEB-2002 (first entry)  
 DT XX  
 DE Propionibacterium acnes immunogenic protein #3739.  
 XX  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO200181581-A2.  
 XX  
 XX 01-NOV-2001.  
 PD XX  
 XX 20-APR-2001; 2001WO-US012865.  
 PF XX  
 XX 21-APR-2000; 2000US-0199047P.  
 PR 02-JUN-2000; 2000US-0208841P.  
 PR 07-JUL-2000; 2000US-0216747P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 XX WPI; 2001-616774/71.  
 DR N-PSDB; AAS959518.  
 DR  
 XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris.  
 PT  
 XX Example 1; SEQ ID NO 4038; 1069pp; English.  
 PS  
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 60 AA;  
 CC  
 CC Query Match 55.6%; Score 5; DB 4; Length 60;  
 CC Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRHP 9  
 Db 13 FLRHP 17  
 RESULT 24  
 ID ABM39362 standard; protein; 60 AA.  
 XX

AC ABM39362;  
 XX 20-OCT-2003 (first entry)  
 DT XX  
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #4038.  
 XX  
 KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
 KW immunostimulant; immune response; vaccine.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO2003033515-A1.  
 XX  
 XX 24-APR-2003.  
 PD XX  
 XX 11-OCT-2002; 2002WO-US032727.  
 PF XX  
 XX 15-OCT-2001; 2001US-00978825.  
 PR XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
 PI Barth B, Vallieue-Douglas J;  
 XX  
 XX WPI; 2003-381789/36.  
 DR N-PSDB; ACF64447.  
 DR  
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a P. acnes protein.  
 PS Example 1; SEQ ID NO 4038; 1481pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
 CC encoding a Propionibacterium acnes protein. The invention also relates to  
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
 CC immunogenic fragments of P. acnes polypeptides. The invention  
 CC additionally encompasses expression vectors and host cells comprising a  
 CC polynucleotide of the invention; antibodies against polypeptides of the  
 CC invention; fusion proteins comprising a polypeptide of the invention; a  
 CC method for stimulating an immune response specific for a P. acnes  
 CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,  
 CC via this method; a vaccine composition (comprising P. acnes polypeptides, or  
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
 CC antigen-presenting cells that express the polypeptide); a method and kit  
 CC for detecting or determining the presence or absence of P. acnes in a  
 CC patient; and a method for inhibiting the development of P. acnes in a  
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
 CC proteins, T cell populations or antigen-presenting cells that express the  
 CC polypeptides are useful for diagnosing, preventing or treating acne  
 CC vulgaris, or for stimulating an immune response specific for a P. acnes  
 CC protein. The polynucleotides can also be used as probes or primers for  
 CC nucleic acid hybridisation. The vaccine composition is useful for the  
 CC stimulation of an immune response against P. acnes, or for treating acne,  
 CC and the kit is useful for performing a diagnostic assay. The present  
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
 CC reading frame) contained within the P. acnes polynucleotides of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 60 AA;  
 CC  
 CC Query Match 55.6%; Score 5; DB 6; Length 60;  
 CC Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRHP 9  
 Db 13 FLRHP 17

RESULT 25  
AAM86926  
ID AAM86926 standard; protein; 62 AA.  
XX AC  
AAM86926;  
XX DT 07-NOV-2001 (first entry)  
XX DE Human immune/haematopoietic antigen SEQ ID NO:14519.  
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX KW cystostatic; gene therapy; vaccine; metastasis.  
XX OS Homo sapiens.  
XX PN WO200157182-A2.  
XX PD 09-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001354.  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236127P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 20-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 08-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 DR WPI; 2001-483426/52.  
 DR N-PSDB; AAK59707.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides;  
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
 XX  
 PS Claim 11; SEQ ID NO 14519; 3071pp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 62 AA;  
 Query Match 55.6%; Score 5; DB 4; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ETWFL 6  
 Db 39 ETWFL 43  
 RESULT 26  
 AAG98737  
 ID AAG98737 standard; protein; 64 AA.  
 XX  
 AC AAG98737;  
 XX  
 DT 21-SEP-2001 (first entry)  
 XX  
 DE Human cell death protective cDNA clone CNI-00720 ORF3 protein, SEQ:267.  
 XX  
 KW Cell death protective; apoptosis; necrosis; human; drug screening;  
 KW cell death-associated disorder; central nervous system disorder;  
 KW psychiatric disorder; neurological disorder; ischaemia-related disorder;  
 KW stroke; cerebral infarction; ischaemic encephalopathy;  
 KW neurodegenerative disorder; Alzheimer's disease; Huntington's disease;  
 KW Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;  
 KW vascular disease; ophthalmological disorder; diabetic retinopathy;  
 KW macular degeneration; hypertension; myocardial infarction;  
 KW atherosclerosis; respiratory disorder; asthma; transgenic animal;  
 KW chronic obstructive pulmonary disease; neoplastic condition; cancer;

KW benign tumour; anaemia; gastrointestinal disorder; gastritis;  
 KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;  
 KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;  
 KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;  
 XX urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.  
 OS Homo sapiens.  
 XX  
 PN WO2000145638-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 11-DEC-2000; 2000WO-US033547.  
 XX  
 PR 14-DEC-1999; 99US-00461697.  
 XX  
 PA (COGE-) COGENT NEUROSCIENCE INC.  
 PI Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;  
 DR WPI; 2001-390297/41.  
 DR N-PSDB; AAH84265, AAH84268.  
 XX  
 PT Novel protective sequence polynucleotides and polypeptides, used to  
 PT identify modulators of their expression and activity, which are used in  
 PT to treat central nervous system conditions, diseases and disorders.  
 XX  
 PS Claim 1; Fig 10C; 325pp; English.  
 XX  
 CC Sequences AAH84132-AAH84370 represent human nucleic acid sequences which  
 CC protect against cell death (i.e., apoptosis or necrosis). Sequences  
 CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,  
 CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,  
 CC while the remaining nucleic acid sequences within the range given above  
 CC represent the open reading frames (ORFs) of these cDNA clones. Sequences  
 CC AAG98610-AAG98829 represent the polypeptides encoded by the cell death  
 CC protective ORFs. The cell death protective cDNA clones are able to  
 CC prevent, delay or reverse progression through the apoptotic or necrotic  
 CC pathways when injected into a cell predisposed to or undergoing cell  
 CC death. The cell death protective nucleic acids and polypeptides can be  
 CC used in the diagnosis and treatment of disorders associated with cell  
 CC death, and to screen for compounds which modulate their activity or  
 CC expression. Such modulators, preferably a small organic molecule, an  
 CC antibody, a ribozyme, or an antisense molecule, can also be used to treat  
 CC cell death-related diseases. Such diseases include those associated with  
 CC the central nervous system including psychiatric or neurological  
 CC disorders, especially ischaemia-related conditions such as strokes, and  
 CC also includes neurodegenerative disorders such as Alzheimer's disease,  
 CC Huntington's disease, or Parkinson's disease. The modulators may also be  
 CC used to treat infections such as meningitis, malaria, or trypanosomiasis;  
 CC vascular diseases such as ischaemic encephalopathy or cerebral infarction;  
 CC eye conditions such as diabetic retinopathy or macular degeneration;  
 CC hypertension; myocardial infarction; atherosclerosis; respiratory  
 CC conditions such as asthma or chronic obstructive pulmonary disease;  
 CC neoplastic conditions such as cancers or benign tumours; blood cell  
 CC conditions such as anaemia; gastrointestinal conditions such as gastritis  
 CC or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney  
 CC disorders such as glomerulonephritis; cystitis; endometriosis; endocrine  
 CC disorders such as Grave's disease or Hashimoto's thyroiditis; skin  
 CC conditions such as dermatitis or urticaria; or immune system disorders  
 CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may  
 CC additionally be used to generate animal models of cell death-associated  
 CC disorders. The present sequence represents a cell death protective  
 CC polypeptide  
 XX  
 SQ Sequence 64 AA;  
 Query Match 55.6%; Score 5; DB 4; Length 64;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRH 8  
 |||||

Db 24 WFLRH 28

RESULT 27  
AAU50032  
ID AAU50032 standard; protein; 64 AA.  
AC AAU50032;  
XX  
DT 27-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #10928.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 20-APR-2001; 2001WO-US012865.  
XX  
PR 21-APR-2000; 2000US-0199047P.  
PR 02-JUN-2000; 2000US-0208841P.  
PR 07-JUL-2000; 2000US-0216747P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI; 2001-616774/71.  
DR N-PSDB; AAS59546.  
XX  
PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.  
XX  
PS Example 1; SEQ ID NO 11227; 1069pp; English.  
XX  
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 64 AA;  
Query Match 55.6%; Score 5; DB 4; Length 64;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 30 FLRHP 34

RESULT 28  
ABM46551  
ID ABM46551 standard; protein; 64 AA.  
XX  
AC ABM46551;  
XX  
DT 20-OCT-2003 (first entry)  
XX  
DE Propionibacterium acnes predicted ORF-encoded polypeptide #11227.  
XX  
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
KW immunostimulant; immune response; vaccine.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO2003033515-A1.  
XX  
PD 24-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032727.  
XX  
PR 15-OCT-2001; 2001US-00978825.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
PI Barth B, Valliave-Dougllass J;  
XX  
DR WPI; 2003-381789/36.  
DR N-PSDB; ACF64475.  
XX  
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.  
XX  
PS Example 1; SEQ ID NO 11227; 1481pp; English.  
XX  
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression vectors and host cells comprising a  
CC polynucleotide of the invention; antibodies against polypeptides of the  
CC invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide); a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the development of P. acnes in a  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC proteins, T cell populations or antigen-presenting cells that express the  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 64 AA;  
Query Match 55.6%; Score 5; DB 6; Length 64;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

CC	format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SO	Sequence 73 AA;
	Query Match 55.6%; Score 5; DB 5; Length 73;
	Best Local Similarity 100.0%; Pred. No. 1.6e+02;
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	5 FLRHP 9
Db	3 FLRHP 7
RESULT 30	
AAG98736	
ID	AAG98736 standard; protein; 75 AA.
XX	
AC	AAG98736;
XX	
DT	21-SEP-2001 (first entry)
XX	
DE	Human cell death protective cDNA clone CNI-00720 ORF2 protein, SEQ:265.
XX	
KW	Cell death protective; apoptosis; necrosis; human; drug screening;
KW	cell death-associated disorder; central nervous system disorder;
KW	psychiatric disorder; neurological disorder; ischaemia-related disorder;
KW	stroke; cerebral infarction; ischaemic encephalopathy;
KW	neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
KW	Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
KW	vascular disease; ophthalmological disorder; diabetic retinopathy;
KW	macular degeneration; hypertension; myocardial infarction;
KW	atherosclerosis; respiratory disorder; asthma; transgenic animal;
KW	chronic obstructive pulmonary disease; neoplastic condition; cancer;
KW	benign tumour; anaemia; gastrointestinal disorder; gastritis;
KW	ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KW	glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KW	Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KW	urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.
XX	
OS	Homo sapiens.
XX	
PN	WO200145638-A2.
XX	
PD	28-JUN-2001.
XX	
PF	11-DEC-2000; 2000WO-US033547.
XX	
PR	14-DEC-1999; 99US-00461697.
XX	
XX	(COGE-) COGENT NEUROSCIENCE INC.
XX	
PI	Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;
XX	
DR	WPI; 2001-390297/41.
DR	N-PSDB; AAH84265, AAH84267.
XX	
PT	Novel protective sequence polynucleotides and polypeptides, used to
PT	identify modulators of their expression and activity, which are used in
PT	to treat central nervous system conditions, diseases and disorders.
XX	
PS	Claim 1; Fig 10B; 325pp; English.
XX	
CC	Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
CC	protect against cell death (i.e. apoptosis or necrosis). Sequences
CC	AAH84132, AAH84134, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
CC	AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
CC	while the remaining nucleic acid sequences within the range given above
CC	represent the open reading frames (ORFs) of these cDNA clones. Sequences
CC	AAH84132-AAH84280 represent the polypeptides encoded by the cell death
CC	protective ORFs. The cell death protective cDNA clones are able to
CC	prevent, delay or reverse progression through the apoptotic or necrotic
CC	pathways, when injected into a cell predisposed to or undergoing cell
CC	death. The cell death protective nucleic acids and polypeptides can be

Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	5 FLRHP 9								
Db	30 FLRHP 34								
RESULT 29									
ABP03674									
ID	ABP03674 standard; protein; 73 AA.								
XX									
AC	ABP03674;								
XX									
DT	25-JUN-2002 (first entry)								
XX									
DE	Human ORFX protein sequence SEQ ID NO:7330.								
XX									
KW	Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;								
KW	hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;								
KW	degenerative disorder; osteoarthritis; neurodegenerative disorder;								
KW	cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;								
KW	hypertension; hypothyroidism; cholesterol ester storage disease;								
KW	immune deficiency; immune disorder; infectious disease;								
KW	autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;								
KW	myasthenia gravis.								
XX									
OS	Homo sapiens.								
XX									
PN	WO200192523-A2.								
XX									
PD	06-DEC-2001.								
XX									
PF	29-MAY-2001; 2001WO-US010836.								
XX									
XX	30-MAY-2000; 2000US-0206132P.								
PR	29-AUG-2000; 2000US-0228716P.								
XX									
XX	(CURA-) CURAGEN CORP.								
PA									
XX									
PI	Shimkets RA, Leach MD;								
XX									
XX	WPI; 2002-106308/14.								
DR	N-PSDB; ABN19426.								
XX									
PT	Novel human polypeptides and polynucleotides useful for diagnosing,								
PT	preventing and treating cardiovascular disease, neurodegenerative,								
PT	hyperproliferative disorders and autoimmune disorders.								
XX									
XX	Disclosure; SEQ ID NO 7330; 1037pp; English.								
PS									
XX									
CC	The present invention describes substantially purified human proteins								
CC	(referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1								
CC	in the specification). ABN15762 to ABN27252 encode the human ORFX								
CC	proteins given in ABP00010 to ABP11500. ORFX proteins are useful for								
CC	treating or preventing a pathology associated with an ORFX-associated								
CC	disorder in humans, and in the manufacture of a medicament for treating a								
CC	syndrome associated with ORFX-associated disorder. ORFX polynucleotide								
CC	sequences can be used in gene therapy. ORFX sequences can be used in the								
CC	treatment of cancer, hyperproliferative disorders, cirrhosis of liver,								
CC	psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,								
CC	osteoarthritis, neurodegenerative disorders, disorders related to organ								
CC	transplantation, cardiovascular diseases, diabetes mellitus, systemic								
CC	lupus erythematosus, hypertension, hypothyroidism, cholesterol ester								
CC	storage disease, various immune deficiencies and disorders, infectious								
CC	diseases, autoimmune disorders such as multiple sclerosis, rheumatoid								
CC	arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host								
CC	disease and autoimmune inflammatory eye disease. ORFX proteins are also								
CC	useful for treating burns, incisions, ulcers, for treating osteoporosis,								
CC	bone degenerative disorders, or periodontal disease, and for gut								
CC	protection or regeneration and treatment of lung or liver fibrosis,								
CC	reperfusion injury in various tissues and conditions resulting from								
CC	systemic cytokine damage. N.B. The sequence data for this patent did not								
CC	form part of the printed specification, but was obtained in electronic								

CC used in the diagnosis and treatment of disorders associated with cell  
 CC death, and to screen for compounds which modulate their activity or  
 CC expression. Such modulators, preferably a small organic molecule, an  
 CC antibody, a ribozyme, or an antisense molecule, can also be used to treat  
 CC cell death-related diseases. Such diseases include those associated with  
 CC the central nervous system including psychiatric or neurological  
 CC disorders, especially ischemia-related conditions such as strokes, and  
 CC also includes neurodegenerative disorders such as Alzheimer's disease,  
 CC Huntington's disease, or Parkinson's disease. The modulators may also be  
 CC used to treat infections such as meningitis, malaria, or trypanosomiasis;  
 CC vascular diseases such as ischaemic encephalopathy or cerebral infarction;  
 CC eye conditions such as diabetic retinopathy or macular degeneration;  
 CC hypertension; myocardial infarction; atherosclerosis; respiratory  
 CC conditions such as asthma or chronic obstructive pulmonary disease;  
 CC neoplastic conditions such as cancers or benign tumours; blood cell  
 CC conditions such as anaemia; gastrointestinal conditions such as gastritis  
 CC or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney  
 CC disorders such as glomerulonephritis; cystitis; endometriosis; endocrine  
 CC disorders such as Grave's disease or Hashimoto's thyroiditis; skin  
 CC conditions such as dermatitis or urticaria; or immune system disorders  
 CC such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may  
 CC additionally be used to generate animal models of cell death-associated  
 CC disorders. The present sequence represents a cell death protective  
 CC polypeptide

SQ Sequence 75 AA;

Query Match 55.6%; Score 5; DB 4; Length 75;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
 DB 35 WFLRH 39

RESULT 31

AM99844  
 ID AM99844 standard; protein; 76 AA.

AC AM99844;

DT 07-JAN-2002 (first entry)

DE Human excretory related polypeptide SEQ ID NO 581.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
 KW antiparkinsonian; antisickling; antianemic; antiarthritic; cancer;  
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; cerebroprotective; anticonvulsant; antifungal;  
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine;  
 KW excretory system.

XX Homo sapiens.

XX WO20015313-A2.

PN 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US001323.

PF 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214866P.

PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226688P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 21-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0239355P.  
 PR 13-OCT-2000; 2000US-0239357P.  
 PR 20-OCT-2000; 2000US-0239379P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.



CC prevention of: (a) cancer, e.g. breast and ovarian cancer and other  
 CC cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal  
 CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's  
 CC disease, allergies, autoimmune haemolytic anaemia, autoimmune  
 CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,  
 CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders  
 CC such as myocardial ischaemia; (d) wound healing; (e) neurological  
 CC diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases  
 CC such as viral, bacterial, fungal and parasitic infections. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 76 AA;

Query Match 55.6%; Score 5; DB 4; Length 76;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0;  
 Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 3 TWFLR 7  
 DB 4 TWFLR 8

RESULT 32  
 AAM42659  
 ID AAM42659 standard; protein; 76 AA.

XX AC AAM42659;

XX DT 22-OCT-2001 (first entry)

DE Human kidney related polypeptide SEQ ID NO 528.

XX Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;  
 KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
 KW neotropic; neuroprotective; antibacterial; virucide; fungicide;  
 KW ophthalmological; antiallergic; hepatotropic; antidiabetic;  
 KW antinflammatory; antitumor; antitumor; antitumor; antitumor;  
 KW gene therapy; cancer; immune disease; infection.

XX OS Homo sapiens.

XX PN WO200155323-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001343.  
 XX 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214986P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 11-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 14-JUL-2000; 2000US-0217496P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.

PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465569/50.

XX N-PSDB; AAI98817.

XX Isolated nucleic acid molecule encoding excretory system antigen is used  
 PT in preventing, treating or ameliorating a medical condition.

XX Claim 11; SEQ ID NO 581; 574pp + Sequence Listing; English.

XX The invention relates to novel excretory system related human  
 CC polynucleotides (AAI98567-AAI99503) and the encoded proteins (AAM99594-  
 CC AAM99913) useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy, especially disorders related  
 CC to the excretory system. The genes are isolated from a range of human  
 CC tissues disclosed in the specification. The nucleic acids, proteins, and  
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment and



PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226273P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 03-DEC-2000; 2000US-0251030P.  
PR 03-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488784/53.

N-PSDB; AA163213.

New isolated nucleic acids and polypeptides, useful for diagnosing, treating and/or preventing human diseases and disorders.

Claim 11; SEQ ID NO 528; 564pp + Sequence Listing; English.

The invention relates to novel kidney related polynucleotides (AA162971-AA163793) and the encoded polypeptides (AA42417-AA42691) collectively known as kidney antigens and the use of such kidney antigens for detecting disorders of the kidney, especially kidney cancer and kidney cancer metastases. The polynucleotides and proteins are also useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at

```

CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 76 AA;
  Query Match      55.6%; Score 5; DB 4; Length 76;
  Best Local Similarity 100.0%; Pred. No. 1.7e+02;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7
   |||||
Db 4 TWFLR 8

RESULT 33
ID ABP34862 standard; protein; 82 AA.
XX
AC ABP34862;
XX
DT 08-JUL-2002 (first entry)
XX
DE Human ORF3835 protein, SEQ ID NO:7670.
XX
KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotrophic; antipsoptic; antidiabetic; cytostatic; neurotic;
KW neuroprotective; antithrombotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS Homo sapiens.
XX
PN WO200190366-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US017076.
XX
PR 24-MAY-2000; 2000US-0206690P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Leach MD, Shinkets RA;
XX
DR WPI; 2002-106200/14.
DR N-PSDB; ABN78888.
XX
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and disorders related to organ
XX transplantation.
XX
PS Claim 10; Page 2162; 2508pp; English.
XX
CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
  cell differentiation, immune modulation, haematopoiesis regulation,
  tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
  chemokinetic activity, haemostatic activity, thrombolytic activity,
  receptor/ligand, antiinflammatory activity, tumour inhibition activity,
  and antifertility activity, and may also be involved in the determination
  of bodily characteristics, fertility and behaviour. ORFX proteins,
  nucleic acids and antibodies may be used in the treatment of cancers,
  other proliferative disorders such as psoriasis and benign tumours,
  neurological disorders such as epilepsy and Alzheimer's disease,
  cardiovascular diseases, immune system disorders, disorders related to
  organ transplantation, disorders of tissue growth and regeneration,
  diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
  storage disease, and infectious diseases caused by viral, bacterial,
  fungal and other pathogens. ORFX nucleic acids may also be used as a
  source of primers and probes, in the detection of ORFX genomic sequences
  or transcripts, in the identification and cloning of homologous
  sequences, in genetic diagnosis, and in forensic biology. The ORFX
  nucleic acids may additionally be used to produce transgenic animals
  which may be useful for studying the function and/or activity of ORFX
  protein, and in drug screening. The ORFX proteins may also be used as
  immunogens to generate specific antibodies, which are useful in the
  diagnosis, treatment and monitoring of ORFX-associated diseases
  XX
  SQ Sequence 82 AA;
    Query Match      55.6%; Score 5; DB 5; Length 82;
    Best Local Similarity 100.0%; Pred. No. 1.8e+02;
    Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7
   |||||
Db 76 TWFLR 80

RESULT 34
ADT58131
ID ADT58131 standard; protein; 85 AA.
XX
AC ADT58131;
XX
DT 13-JAN-2005 (first entry)
XX
DE Plant polypeptide, SEQ ID 8208.
XX
KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
OS Viridiplantae.
XX
XX US2004216190-A1.
XX
XX 28-OCT-2004.
XX
XX 18-DEC-2003; 2003US-00739930.
XX
XX 28-APR-2003; 2003US-00424599.
XX
XX 28-APR-2003; 2003US-00425115.
XX
XX (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
XX
XX WPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
XX genetics, and in particular for producing transgenic plants with improved
XX biological characteristics.
XX
XX Claim 2; SEQ ID NO 8208; 14pp; English.
XX

```



XX  
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
PI Barth B, Vallieue-Doughlass J;  
XX  
XX WPI; 2003-381789/36.  
DR N-PSDB; ACF64480.  
XX  
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.  
XX  
XX Example 1; SEQ ID NO 12713; 1481pp; English.  
XX  
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression vectors and host cells comprising a  
CC polynucleotide of the invention; antibodies against polypeptides of the  
CC invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide); a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the development of P. acnes in a  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC proteins, T cell populations or antigen-presenting cells that express the  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 86 AA;  
SQ  
Query Match 55.6%; Score 5; DB 6; Length 86;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FLRHP 9  
Db 2 FLRHP 6  
RESULT 37  
ADX94950  
ID ADX94950 standard; protein; 89 AA.  
XX  
XX ADX94950;  
XX  
XX 21-APR-2005 (first entry)  
XX Plant full length insert polypeptide seqid 57614.  
XX  
XX plant protectant; plant growth regulant; gene therapy; plant;  
KW recombinant DNA construct; physical array; plant breeding marker;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
KW extreme osmotic condition; pathogen tolerance; pest tolerance;  
KW growth rate; cell cycle pathway; disease resistance;  
KW galactomannan production; lignin production; plant growth regulator;  
KW yield; plant growth; plant development; seed oil; protein yield;  
KW protein content.  
XX  
XX Unidentified.  
OS

XX  
PN US2004034888-A1.  
XX  
XX 19-FEB-2004.  
XX  
XX 28-APR-2003; 2003US-00425114.  
XX  
XX 06-MAY-1999; 99US-00304517.  
XX 05-NOV-2001; 2001US-00985678.  
XX  
XX (LIUJ/) LIU J.  
XX (ZHOU/) ZHOU Y.  
XX (KOVA/) KOVALIC D K.  
XX (SCRE/) SCREEN S E.  
XX (TABAS/) TABASKA J E.  
XX (CAOY/) CAO Y.  
XX  
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;  
PI WPI; 2004-180133/17.  
XX  
XX New recombinant DNA construct, useful for improving plant tolerance to  
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
PT pests, for conferring increased resistance to plant disease, or for  
PT improving yield.  
XX  
XX Claim 1; SEQ ID NO 57614; 15pp; English.  
XX  
XX The invention describes a recombinant DNA construct comprising a  
CC polynucleotide consisting of a sequence encoding an amino acid sequence  
CC available in electronic form from the US patent office at  
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide  
CC of the invention are also useful in physical arrays of molecules and as  
CC plant breeding markers. The recombinant DNA construct is useful for  
CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
CC plant cells by modification of the cell cycle pathway, for conferring  
CC increased resistance to plant disease, for producing galactomannan,  
CC lignin or plant growth regulators, for increasing the rate of homologous  
CC recombination in plants, for improving yield by modification of  
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
CC or by providing improved plant growth and development under at least one  
CC stress condition or for modifying seed oil or protein yield and/or  
CC content. This is the amino acid sequence of a plant full length insert  
CC polypeptide that can be used in the recombinant DNA construct of the  
CC invention.  
XX  
XX Sequence 89 AA;  
SQ  
Query Match 55.6%; Score 5; DB 8; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 WFLRH 8  
Db 6 WFLRH 10  
RESULT 38  
ABM94143  
ID ABM94143 standard; protein; 98 AA.  
XX  
XX ABM94143;  
XX  
XX 02-JUN-2005 (first entry)  
XX  
XX M. xanthus protein sequence, seq id 13342.  
DE  
XX Transgenic plant; DNA replication; gene regulation; gene expression.  
KW  
XX Myxococcus xanthus.  
OS  
XX  
XX US6833447-B1.  
PN

```

XX PD 21-DEC-2004.
XX PF 10-JUL-2001; 2001US-00902540.
XX PF 10-JUL-2000; 2000US-0217883P.
XX PA (MONS ) MONSANTO TECHNOLOGY LLC.
XX PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX PT New substantially purified Myxococcus xanthus nucleic acid molecule
XX PT encoding a nitrite reductase, useful for determining gene expression,
XX PT identifying mutations in a gene of interest, and for constructing
XX PT mutations in a gene of interest.
XX PS Example 2; SEQ ID NO 13342; 25pp; English.
XX CC The invention relates to a substantially purified nucleic acid molecule
XX CC encoding a nitrite reductase of SEQ ID NO. 11976. Further disclosed is a
XX CC recombinant DNA construct for expression of a nitrite reductase gene in a
XX CC plant cell, and a plant cell comprising the recombinant DNA construct.
XX CC The nucleic acid is useful for determining gene expression, identifying
XX CC mutations in a gene of interest, and for constructing mutations in a gene
XX CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
XX CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from USPTO
XX SQ Sequence 98 AA;

Query Match 55.6%; Score 5; DB 9; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9
Db 29 FLRHP 33

RESULT 39
ADCL4235
ID ADC14235 standard; protein; 102 AA.
AC ADC14235;
XX 18-DEC-2003 (first entry)
XX Human enzyme ENZM-41.
XX enzyme; human; ENZM; cytostatic; antiarteriosclerotic; antidiabetic;
XX anticonvulsant; nootropic; neuroprotective; cerebroprotective; anti-HIV;
XX anti-allergic; anti-inflammatory; thyromimetic; gene therapy;
XX cell proliferative disorder; endocrine disorder; neurological disorder;
XX immune system disorder; inflammatory disorder; developmental disorder;
XX reproductive disorder; vesicle-trafficking disorder; infection.
XX OS Homo sapiens.
XX WO2003042357-A2.
XX 22-MAY-2003.
XX 26-SEP-2002; 2002WO-US031096.
XX 28-SEP-2001; 2001US-0326388P.
XX 12-OCT-2001; 2001US-0328979P.
XX 19-OCT-2001; 2001US-0346034P.
XX 26-OCT-2001; 2001US-0348284P.
XX 08-NOV-2001; 2001US-0338048P.
XX 16-NOV-2001; 2001US-0332340P.

```

```

PR 14-DEC-2001; 2001US-0340357P.
PR 29-MAR-2002; 2002US-0368722P.
PR 29-MAR-2002; 2002US-0368799P.
PR 17-MAY-2002; 2002US-0381558P.
PR 07-JUN-2002; 2002US-0387119P.
PR 21-JUN-2002; 2002US-0390662P.
XX (INCY-) INCYTE GENOMICS INC.
XX Yang J, Lu DAM, Yue H, Elliott VS, Warren BA, Duggan BM;
XX Forsythe IJ, Lee EA, Hafalia AJA, Ramkumar J, Chawla NK, Baughn MR;
XX Becha SD, Gorvad AE, Tran UK, Li JX, Yao MG, Ison CH, Griffin JA;
XX Lee SY, Chang H, Emerling BM, Tang YT, Lal PG, Kable AE;
XX Marquis JP, Jiang X, Jackson AA, Zebartadian Y, Swarnakar A;
XX Wilson AD, Jin P, Richardson TW, Bhatia U, Burrill JD, Lee S;
XX Blake JJ, Ho A, Zheng W, Gao J;
XX WPI; 2003-449567/42.
XX N-PSDB; ADC14288.
XX New human enzymes (ENZM), useful for diagnosing, treating and preventing
XX diseases or conditions associated with the aberrant ENZM expression e.g.
XX cancer, diabetes, epilepsy, or infections.
XX Claim 1; SEQ ID NO 41; 416pp; English.
XX The invention relates to a novel isolated human enzyme (ENZM)
XX polypeptide. A polypeptide of the invention has cytostatic,
XX antiarteriosclerotic, antidiabetic, anticonvulsant, nootropic,
XX neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
XX anti-inflammatory, and thyromimetic activity. A polynucleotide encoding a
XX polypeptide of the invention may have a use in gene therapy. The
XX polypeptides and polynucleotides are useful in diagnosing, treating and
XX preventing diseases or conditions associated with the decreased
XX expression or overexpression of ENZM, such as cell proliferative (e.g.
XX cancer, atherosclerosis), endocrine (e.g. diabetes), neurological (e.g.
XX epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
XX allergies), developmental (e.g. Hypothyroidism, Cushing's syndrome),
XX reproductive and vesicle-trafficking disorders, or infections. These are
XX also useful in assessing the effects of exogenous compounds on the
XX expression of nucleic acid and amino acid sequences of ENZM. The ENZM or
XX its fragments are useful in screening compounds for effectiveness as
XX agonist or antagonist of the polypeptides, or in altering the expression
XX of the target polynucleotide and compounds that specifically bind to or
XX modulate the activity of the polypeptide. The microarray is useful in
XX monitoring or measuring protein-protein interactions, drug-target
XX interactions, and gene expression profiles. The sequences shown in
XX CC ADC14195-ADC14247 represent ENZM proteins of the invention.
XX SQ Sequence 102 AA;

Query Match 55.6%; Score 5; DB 7; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9
Db 57 FLRHP 61

RESULT 40
AAV74113
ID AAY74113 standard; protein; 104 AA.
XX AAY74113;
XX 14-MAR-2000 (first entry)
XX Human prostate tumor EST fragment derived protein #300.
XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
XX treatment.
XX

```

```

OS Homo sapiens.
XX
XX DE19820190-A1.
XX
XX 04-NOV-1999.
XX
XX 28-APR-1998; 98DE-01020190.
XX
XX 28-APR-1998; 98DE-01020190.
XX
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-621386/54.
XX N-PSDB; AA252956.
XX
XX New human nucleic acid sequences from pancreatic tumors, and related
XX proteins.
XX
XX Claim 23; Page 434; 502pp; German.
XX
XX This invention describes novel polypeptides and their encoding nucleic
XX acids derived from human pancreatic tumor tissue which have cytostatic
XX activity. The sequences are also useful in producing pharmaceutical
XX compositions for treatment of pancreatic tumors. AA73814-Y74252
XX represent protein fragments encoded by the human pancreatic tumor cDNA
XX library derived expressed sequence tag (EST) sequences represented in
XX CC AA252858-Z53014
XX
XX Sequence 104 AA;
SQ
Query Match 55.6%; Score 5; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 FLRHP 9
DB 85 FLRHP 89
Search completed: August 31, 2006, 10:46:44
Job time : 116.75 secs

```



103 4 44.4 115 2 G72568 hypothetical prote  
104 4 44.4 115 2 F72779 hypothetical prote  
105 4 44.4 116 2 H95414 hypothetical prote  
106 4 44.4 118 2 G95121 Tn5252, Orf 10 pro  
107 4 44.4 120 2 C86882 hypothetical prote  
108 4 44.4 121 2 S76514 hypothetical prote  
109 4 44.4 121 2 T08717 hypothetical prote  
110 4 44.4 122 2 D72756 hypothetical prote  
111 4 44.4 123 2 H90236 conserved hypothet  
112 4 44.4 123 2 S09822 hypothetical prote  
113 4 44.4 123 2 AH2707 conserved hypothet  
114 4 44.4 124 2 S03521 Ig kappa chain pre  
115 4 44.4 124 2 C75359 hypothetical prote  
116 4 44.4 125 2 S07739 hypothetical prote  
117 4 44.4 125 2 H81890 probable phage rep  
118 4 44.4 125 2 F82834 hypothetical prote  
119 4 44.4 128 2 S76468 hypothetical prote  
120 4 44.4 128 2 AF2143 hypothetical prote  
121 4 44.4 129 2 A75346 hypothetical prote  
122 4 44.4 130 2 A41911 oxoglutarate dehyd  
123 4 44.4 132 2 T50389 homolog to yeast PK  
124 4 44.4 133 2 B44370 probable G-protein  
125 4 44.4 134 2 F85589 hypothetical prote  
126 4 44.4 134 2 D90739 hypothetical prote  
127 4 44.4 134 2 F64817 probable membrane  
128 4 44.4 134 2 C84385 hypothetical prote  
129 4 44.4 135 2 S31682 inhibin beta-A cha  
130 4 44.4 136 2 T18052 DEAH box protein a  
131 4 44.4 136 2 H72633 hypothetical prote  
132 4 44.4 137 2 F29380 Ig heavy chain pre  
133 4 44.4 137 2 A71308 hypothetical prote  
134 4 44.4 137 2 A75292 hypothetical prote  
135 4 44.4 137 2 S40760 hypothetical prote  
136 4 44.4 141 1 HALZC hemoglobin alpha-1  
137 4 44.4 141 2 T06224 hypothetical prote  
138 4 44.4 141 2 F46427 probable transcrip  
139 4 44.4 142 2 F82239 conserved hypothet  
140 4 44.4 143 2 E89889 probable membrane  
141 4 44.4 143 2 A10838 trans-regulatory s  
142 4 44.4 144 1 F46335 hypothetical prote  
143 4 44.4 146 2 S34584 hypothetical prote  
144 4 44.4 146 2 B71430 hypothetical prote  
145 4 44.4 147 2 C90094 hypothetical prote  
146 4 44.4 149 2 S48927 hypothetical prote  
147 4 44.4 150 2 T03586 glycine-rich RNA-b  
148 4 44.4 150 2 S37009 transposase (clone  
149 4 44.4 150 2 S37011 transposase (clone  
150 4 44.4 150 2 C86754 prophage p12 prote

## ALIGNMENTS

RESULT 1  
A03862  
hypothetical protein E-115 - human adenovirus 2  
C:Species: Mastadenovirus h2 (human adenovirus 2)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C:Accession: A03862  
R:Gingeras, T.R.; Sciaky, D.; Gelinias, R.E.; Bing-Dong, J.; Yen, C.E.; Kelly, M.M.; Bull  
J. Biol. Chem. 257, 13475-13491, 1982  
A:Title: Nucleotide sequences from the adenovirus-2 genome.  
A:Reference number: A92351; MUID:83056843; PMID:7142161  
A:Accession: A03862  
A:Molecule type: DNA  
A:Residues: 1-115 <GIN>  
A:Cross-references: UNIPROT:P03290; UNIPARC:UPI00001392B5

Query Match 55.6%; Score 5; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2 ETWFL 6

Db 3 ETWFL 7  
RESULT 2  
S14077  
Ig kappa chain - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999  
C:Accession: S14077  
R:Schwager, J.; Buerckert, N.; Schwager, M.; Wilson, M.  
EMBO J. 10, 505-511, 1991  
A:Title: Evolution of immunoglobulin light chain genes: analysis of Xenopus Igl isotypes  
A:Reference number: S14076; MUID:91160503; PMID:1705882  
A:Accession: S14077  
A:Molecule type: mRNA  
A:Residues: 1-132 <SCH>  
A:Cross-references: UNIPARC:UPI000017698D  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 55.6%; Score 5; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
Db 48 TWFLR 52

RESULT 3  
B88102  
protein W09G10.5 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: B88102  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_eleg  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: B88102  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-168 <STO>  
A:Cross-references: UNIPROT:O16641; UNIPARC:UPI0000075172; GB:chr\_II; PIDN:AB66113.1; PII  
C:Gene: W09G10.5  
C:Genetics:  
A:Map position: 2  
A:Superfamily: Caenorhabditis elegans hypothetical protein C31G12.2

Query Match 55.6%; Score 5; DB 2; Length 168;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7  
Db 30 TWFLR 34

RESULT 4  
H69850  
mutator MutT protein homolog yjHb - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: H69850  
R:Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero  
C.; Bron, S.; Broutillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallier  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;



Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: H69850  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-170 <KUN>  
A;Cross-references: UNIPROT:O34488; UNIPARC:UPI0000060253; GB:Z99110; GB:AL009126; NID:9  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: yjhb

Query Match 55.6%; Score 5; DB 2; Length 170;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FLRHP 9  
Db 159 FLRHP 163

RESULT 5  
QJ0547  
stripe disease-specific protein - rice stripe virus (isolate T)  
C;Species: rice stripe virus  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: QJ0547  
R;Zhu, Y.; Hayakawa, T.; Toriyama, S.  
J. Gen. Virol. 73, 1309-1312, 1992  
A;Title: Complete nucleotide sequence of RNA 4 of rice stripe virus isolate T, and complete nucleotide sequence of the complementary DNA of the same isolate.  
A;Reference number: QJ0547  
A;Accession: QJ0547  
A;Molecule type: genomic RNA  
A;Residues: 1-178 <ZHU>  
A;Cross-references: UNIPROT:Q00844; UNIPARC:UPI00000028CF; GB:D10979; DDBJ:D01164; NID:9  
C;Superfamily: maize stripe virus major noncapsid protein

Query Match 55.6%; Score 5; DB 2; Length 178;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FLRHP 9  
Db 123 FLRHP 127

RESULT 6  
F90534  
transcription antitermination protein [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C;Species: Mycoplasma pulmonis  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C;Accession: F90534  
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, I.; Nucleic Acids Res. 29, 2145-2153, 2001  
A;Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*.  
A;Reference number: A99512; MUID:21267165; PMID:11353084  
A;Accession: F90534  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-196 <KUR>  
A;Cross-references: UNIPROT:Q98R28; UNIPARC:UPI0000004589; GB:AL445566; PID:G14089595; E  
A;Experimental source: strain UAB CTIP  
C;Genetics:  
A;Gene: MYPV 1820  
A;Genetic code: SGC3

Query Match 55.6%; Score 5; DB 2; Length 196;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 ETWFL 6  
Db 82 ETWFL 86

RESULT 7  
T33341  
hypothetical protein K07D4.5 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 31-Dec-2004  
C;Accession: T33341  
R;Henkhaus, J.; Wohldmann, P.  
submitted to the EMBL Data Library, July 1998  
A;Description: The sequence of C. elegans cosmid K07D4.  
A;Reference number: Z21327  
A;Accession: T33341  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-208 <HEN>  
A;Cross-references: UNIPROT:O76574; UNIPARC:UPI0000007CD7A; EMBL:AF077534; PIDN:AAC26289.1  
A;Experimental source: strain Bristol N2; clone K07D4  
C;Genetics:  
A;Gene: CBSP:K07D4.5  
A;Map position: 2  
A;Introns: 25/3; 68/1; 127/1; 160/2

Query Match 55.6%; Score 5; DB 2; Length 208;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 TWFLR 7  
Db 159 TWFLR 163

RESULT 8  
T35525  
probable two component response regulator - *Streptomyces coelicolor*  
C;Species: *Streptomyces coelicolor*  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
C;Accession: T35525  
R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, March 1999  
A;Reference number: Z21581  
A;Accession: T35525  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-221 <SEE>  
A;Cross-references: UNIPROT:Q9X802; UNIPARC:UPI000000DAF6B; EMBL:AL049497; PIDN:CAB39870.1  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Superfamily: ompR protein; response regulator homology

Query Match 55.6%; Score 5; DB 2; Length 221;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 FLRHP 9  
Db 163 FLRHP 167

RESULT 9  
B87657  
conserved hypothetical protein CC3292 [imported] - *Caulobacter crescentus*  
C;Species: *Caulobacter crescentus*  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: B87657

R.;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
N., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: AB7249; MUID:21173698; PMID:11259647  
A:Accession: B87657  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-224 <STO>  
A:Cross-references: UNIPROT:Q9A3B2; UNIPARC:UPI00000C7A0F; GB:AE005673; NID:gi13424986; E  
C:Genetics:  
A:Gene: CC3292

Query Match 55.6%; Score 5; DB 2; Length 224;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
|||||  
DB 97 WFLRH 101

RESULT 10  
C88939  
protein C05E4.8 [imported] - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: C88939  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gcc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: C88939  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-225 <STO>  
A:Cross-references: UNIPROT:O17356; UNIPARC:UPI0000082F2F; GB:chr\_V; PIDN:AAB71277.1; PI  
A:Note: similar to *C. elegans* TC3 transposase (SP:34257)  
C:Genetics:  
A:Gene: C05E4.8  
A:Map position: 5  
C:Superfamily: *Caenorhabditis* transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
DB 184 FLRHP 188

RESULT 11  
AF0656  
conserved hypothetical protein STY1354 [imported] - *Salmonella enterica* subsp. *enterica*  
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
A:Note: this species has also been called *Salmonella typhi*  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AF0656  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AF0656  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-235 <PAR>

A:Cross-references: UNIPARC:UPI0000059F2E; GB:AL513382; PIDN:CAD01623.1; PID:gi6502477; C  
C:Genetics:  
A:Gene: STY1354

Query Match 55.6%; Score 5; DB 2; Length 235;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8  
|||||  
DB 65 WFLRH 69

RESULT 12  
T27636  
hypothetical protein ZC64.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T27636  
R:Bentley, D.  
submitted to the EMBL Data Library, October 1995  
A:Description: The sequence of *C. elegans* cosmid ZC64.  
A:Reference number: Z20397  
A:Accession: T27636  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-241 <BEN>  
A:Cross-references: UNIPROT:Q23379; UNIPARC:UPI0000081BA9; EMBL:U39740; PIDN:AAA80427.1;  
C:Genetics:  
A:Gene: CESP:ZC64.1  
C:Superfamily: *Caenorhabditis* transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 241;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
DB 204 FLRHP 208

RESULT 13  
AI2644  
flagellar basal body rod protein [imported] - *Agrobacterium tumefaciens* (strain C58, Dupa  
C:Species: *Agrobacterium tumefaciens*  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 05-Oct-2004  
C:Accession: AI2644  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AI2644  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-244 <KUR>  
A:Cross-references: UNIPROT:O34170; UNIPARC:UPI00000D1464; GB:AE008688; PIDN:AAL41575.1;  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: flgF  
A:Map position: circular chromosome  
C:Superfamily: rod protein flgF

Query Match 55.6%; Score 5; DB 2; Length 244;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
DB 168 FLRHP 172

```

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-270 <BLAT>
A;Cross-references: UNIPROT:P77147; UNIPARC:UPI000013A9BC; GB:AE000262; GB:U00096; NID:9
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 05-Oct-2004
C;Accession: A97427
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A97427
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-244 <KUR>
A;Cross-references: UNIPROT:O34170; UNIPARC:UPI00000D1464; GB:AE007869; PIDN:AAK86370.1;
C;Genetics:
A;Gene: AGR_C_982
A;Map position: circular chromosome
C;Superfamily: rod protein flgF

Query Match 55.6%; Score 5; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9
Db 168 FLRHP 172

RESULT 15
S75903
hypothetical protein - Synecocystis sp. (strain PCC 6803)
C;Species: Synecocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S75903
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75903
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-247 <KAN>
A;Cross-references: UNIPROT:P74268; UNIPARC:UPI00001290CE; EMBL:D90913; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 55.6%; Score 5; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9
Db 209 FLRHP 213

RESULT 16
E64924
hypothetical protein b1669 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: E64924
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E64924

```

```

A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-270 <BLAT>
A;Cross-references: UNIPROT:P77147; UNIPARC:UPI000013A9BC; GB:AE000262; GB:U00096; NID:9
C;Species: Escherichia coli hypothetical protein b1669
C;Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRH 8
Db 20 WFLRH 24

RESULT 17
D85774
hypothetical protein Z2696 [imported] - Escherichia coli (strain O157:H7, substrain EDL9)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85774
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85774
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <STO>
A;Cross-references: UNIPROT:Q8X618; UNIPARC:UPI00000D0C14; GB:AE005174; NID:GI2515668; P;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z2696
C;Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRH 8
Db 20 WFLRH 24

RESULT 18
H90925
hypothetical protein ECs2376 [imported] - Escherichia coli (strain O157:H7, substrain RIM
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90925
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90925
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <HAY>
A;Cross-references: UNIPROT:Q8X618; UNIPARC:UPI00000D0C14; GB:BA000007; PIDN:BA035799.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2376
C;Superfamily: Escherichia coli hypothetical protein b1669

Query Match 55.6%; Score 5; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRH 8
Db 20 WFLRH 24

```

Db 20 WFLRH 24

RESULT 19  
E91027  
hypothetical protein ECs3189 [imported] - Escherichia coli (strain O157:H7, substrain R1  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 31-Dec-2004  
C:Accession: E91027  
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno  
A:Reference number: A95629, MUID:21156231, PMID:11258796  
A:Accession: E91027  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-296 <HAY>  
A:Cross-references: UNIPROT:Q8XCT0; UNIPARC:UPI00000D0433; GB:BA000007; PIDN:BAB36612.1;  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
C:Genetics:  
C:Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 296;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 17 FLRHP 21

RESULT 20  
F85871  
hypothetical protein yfci [imported] - Escherichia coli (strain O157:H7, substrain EDL93  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Dec-2004  
C:Accession: F85871  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480, MUID:21074935, PMID:11206551  
A:Accession: F85871  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-296 <STO>  
A:Cross-references: UNIPROT:Q8XCT0; UNIPARC:UPI00000D0433; GB:AE005174; NID:g12516661; F  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
C:Gene: yfci  
C:Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 296;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 17 FLRHP 21

RESULT 21  
G65002  
hypothetical protein b2305 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 31-Dec-2004  
C:Accession: G65002  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: G65002  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-296 <BLAT>  
A:Cross-references: UNIPROT:P77768; UNIPARC:UPI0000047C8B; GB:AE000319; GB:U00096; NID:g1  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 296;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 17 FLRHP 21

RESULT 22  
T32681  
hypothetical protein K07C6.14 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
A:Accession: T32681  
R:Wagner-McPherson, C.; Gillam, B.  
submitted to the EMBL Data Library, December 1997  
A:Description: The sequence of C. elegans cosmid K07C6.  
A:Reference number: Z21209  
A:Accession: T32681  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-300 <WAG>  
A:Cross-references: UNIPROT:O44645; UNIPARC:UPI0000079677; EMBL:AF039049; PIDN:AAB94256.1  
A:Experimental source: strain Bristol N2; clone K07C6  
C:Genetics:  
A:Gene: CESP:K07C6.14  
A:Map position: 5  
C:Superfamily: Caenorhabditis transposon Tc1 hypothetical 32K protein

Query Match 55.6%; Score 5; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||  
Db 231 FLRHP 235

RESULT 23  
D64122  
hypothetical protein HI1424 - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
C:Accession: D64122  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: D64122  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-304 <TIGR>  
A:Cross-references: UNIPROT:P45198; UNIPARC:UPI000013AAB2; GB:U32821; GB:LA2023; NID:g15

Query Match 55.6%; Score 5; DB 2; Length 304;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
|||||

Db 60 FLRHP 64

RESULT 24  
AH0966

conserved hypothetical protein STY4020 [imported] - Salmonella enterica subsp. enterica  
C/Species: Salmonella enterica subsp. enterica serovar Typhi  
A/Note: this species has also been called Salmonella typhi  
C/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 31-Dec-2004  
C/Accession: AH0966  
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A/Reference number: AB0502; MUID:21534947; PMID:11677608  
A/Accession: AH0966  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-313 <PAR>  
A/Cross-references: UNIPARC:UPI000005A6D2; GB:AL513382; PIDN:CAD03228.1; PID:g16504856;  
C/Genetics:  
A/Gene: STY4020  
C/Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 313;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 17 FLRHP 21

RESULT 25  
S48036

hypothetical protein - kiwi fruit  
C/Species: Actinidia chinensis var. deliciosa (kiwi fruit)  
C/Date: 26-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 17-Mar-1999  
C/Accession: S48036  
R/Ledger, S.E.; Gardner, R.C.  
Plant Mol. Biol. 25, 877-886, 1994  
A/Title: Cloning and characterization of five cDNAs for genes differentially expressed d  
A/Reference number: S48035; MUID:94355660; PMID:8075403  
A/Accession: S48036  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-317 <LED>  
A/Cross-references: UNIPARC:UPI000012DB29; EMBL:L27809; NID:g450236; PID:g450237  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994

Query Match 55.6%; Score 5; DB 2; Length 317;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 39 FLRHP 43

RESULT 26  
AE1068

probable membrane protein STY4875 [imported] - Salmonella enterica subsp. enterica serov  
C/Species: Salmonella enterica subsp. enterica serovar Typhi  
A/Note: this species has also been called Salmonella typhi  
C/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 31-Dec-2004  
C/Accession: AE1068  
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A/Reference number: AB0502; MUID:21534947; PMID:11677608  
A/Accession: AE1068  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-321 <PAR>  
A/Cross-references: UNIPARC:UPI000005A9BD; GB:AL513382; PIDN:CAD03364.1; PID:g16505636;  
C/Genetics:  
A/Gene: STY4875  
C/Superfamily: human PML-1 protein

Query Match 55.6%; Score 5; DB 2; Length 321;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 17 FLRHP 21

RESULT 27  
H64491

hypothetical protein M1537 - Methanococcus jannaschii  
C/Species: Methanococcus jannaschii  
C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C/Accession: H64491  
R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A/Reference number: A64300; MUID:96337999; PMID:8688087  
A/Accession: H64491  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-343 <BUL>  
A/Cross-references: UNIPROT:Q58932; UNIPARC:UPI000013AD2A; GB:U67594; GB:L77117; NID:g15  
C/Genetics:  
A/Map position: REV1515744-1514713  
C/Superfamily: Methanococcus jannaschii hypothetical protein M1537

Query Match 55.6%; Score 5; DB 2; Length 343;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 105 FLRHP 109

RESULT 28  
A46567

tetracycline resistance protein - Streptomyces rimosus  
C/Species: Streptomyces rimosus  
C/Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 09-Jul-2004  
C/Accession: A46567  
R/Reynes, J.P.; Calmels, T.; Drocourt, D.; Tiraby, G.  
J. Gen. Microbiol. 134, 585-598, 1988  
A/Title: Cloning, expression in Escherichia coli and nucleotide sequence of a tetracycli  
A/Reference number: A46567; MUID:89036114; PMID:3053973  
A/Accession: A46567  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-347 <REY>  
A/Cross-references: UNIPROT:P14551; UNIPARC:UPI0000136B1D; GB:M20370; NID:g153503; PIDN:7

Query Match 55.6%; Score 5; DB 2; Length 347;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TWFLR 7  
Db 11111

```

Db          313 TWFLR 317

RESULT 29
hypothetical protein T25G12.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34382
R:Du, Z.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid T25G12.
A:Reference number: Z21515
A:Accession: T34382
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-358 <DUZ>
A:Cross-references: UNIPROT:Q22789; UNIPARC:UPI0000081C10; EMBL:U43283; PIDN:AAC69023.1
A:Experimental source: strain Bristol N2; clone T25G12
C:Genetics:
A:Gene: CESP:T25G12.9
A:Map position: X
C:Superfamily: Caenorhabditis transposon Tc1 hypothetical 32K protein

Query Match          55.6%; Score 5; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
Db          231 FLRHP 235

RESULT 30
T06460
anthranilate phosphoribosyltransferase (EC 2.4.2.18) - garden pea (fragment)
N:Alternate names: phosphoribosylanthranilate transferase
C:Species: Pisum sativum (garden pea)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06460
R:Sato, N.; Kazuno, A.; Ohta, N.; Ohshima, K.
submitted to the EMBL Data Library, June 1996
A:Description: Isolation of a pea cDNA for phosphoribosylanthranilate transferase.
A:Reference number: Z15694
A:Accession: T06460
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-368 <SAT>
A:Cross-references: UNIPROT:Q43085; UNIPARC:UPI00000A9D1C; EMBL:D86180; PIDN:BAAL3032.1
A:Experimental source: var. Alaska
C:Genetics:
A:Gene: PAT1
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match          55.6%; Score 5; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
Db          339 FLRHP 343

RESULT 31
E64593
2-oxoacid:ferredoxin oxidoreductase (EC 1.2.7.-) alpha chain - Helicobacter pylori (stra
N:Alternate names: 2-oxoacid:ferredoxin oxidoreductase (CoA-acetylating)
C:Species: Helicobacter pylori
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: E64593
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

```

```

Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.A
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:9739467; PMID:9252185
A:Accession: E64593
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-375 <TOM>
A:Cross-references: UNIPROT:Q25311; UNIPARC:UPI00000D30AB; GB:AE000572; GB:AE000511; NID:
C:Superfamily: Helicobacter pylori 2-oxoacid ferredoxin oxidoreductase; 2-oxoacid ferred
C:Keywords: oxidoreductase
F:5-186/Domain: 2-oxoacid ferredoxin oxidoreductase homology <FEO>

Query Match          55.6%; Score 5; DB 1; Length 375;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
Db          124 FLRHP 128

RESULT 32
G71919
chain of 2-oxoglutarate oxidoreductase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: G71919
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: G71919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <ARN>
A:Cross-references: UNIPROT:Q9ZLP1; UNIPARC:UPI00000D364E; GB:AE001486; GB:AE001439; NID:
A:Experimental source: strain J99
C:Genetics:
A:Gene: oorA
C:Superfamily: Helicobacter pylori 2-oxoacid ferredoxin oxidoreductase; 2-oxoacid ferred
F:5-186/Domain: 2-oxoacid ferredoxin oxidoreductase homology <FEO>

Query Match          55.6%; Score 5; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          5 FLRHP 9
          |||||
Db          124 FLRHP 128

RESULT 33
H96773
hypothetical protein FLM20.17 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H96773
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: H96773
A:Status: preliminary

```

A:Molecule type:..DNA  
A:Residues: 1-378 <STO>  
A:Cross-references: UNIPROT:Q9CA65; UNIPARC:UPI00000A3B86; GB:AE005173; NID:g6539251; PID:158168  
C:Genetics:  
A:Gene: FLN20.17  
A:Map position: 1  
C:Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match 55.6%; Score 5; DB 2; Length 378;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 372 FLRHP 376

RESULT 34  
158168  
growth factor arg3.1 - rat  
A:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I58168; I59386  
R:Lylford, G.L.; Yamagata, K.; Kaufmann, W.E.; Barnes, C.A.; Sanders, L.K.; Copeland, N.C.  
Neuron 14, 433-445, 1995  
A:Title: Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-  
A:Reference number: I58168; MUID:95161073; PMID:7857651  
A:Accession: I58168  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-396 <RES>  
A:Cross-references: UNIPROT:Q62743; UNIPARC:UPI00000E5C7E; EMBL:U19866; NID:g644828; PID:158168  
R:Link, W.; Konietzko, U.; Kauselmann, G.; Krug, M.; Schwanke, B.; Frey, U.; Kuhl, D.  
Proc. Natl. Acad. Sci. U.S.A. 92, 5734-5738, 1995  
A:Title: Somatodendritic expression of an immediate early gene is regulated by synaptic  
A:Reference number: I59386; MUID:95296386; PMID:7777577  
A:Accession: I59386  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-208, 'V', 210-396 <RE2>  
A:Cross-references: UNIPARC:UPI00000E79AC; EMBL:Z46925; NID:g854413; PID:CAA87033.1; PID:158168  
C:Genetics:  
A:Gene: Arc  
C:Superfamily: rat growth factor arg3.1

Query Match 55.6%; Score 5; DB 2; Length 396;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9  
Db 336 FLRHP 340

RESULT 35  
AC3235  
nitrotriacetate monooxygenase nrtA [imported] - Agrobacterium tumefaciens (strain C58,  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AC3235  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-415 <KUR>  
A:Cross-references: UNIPROT:Q8U674; UNIPARC:UPI00000D276A; GB:AE008690; PID:NAL46297.1;

A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: nrtA  
A:Genome: plasmid  
C:Superfamily: nitrotriacetate monooxygenase

Query Match 55.6%; Score 5; DB 2; Length 415;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ETWFL 6  
Db 348 ETWFL 352

RESULT 36  
B86434  
protein T17H7.13 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: B86434  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: B86434  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-415 <STO>  
A:Cross-references: UNIPROT:Q9SY27; UNIPARC:UPI00000A99B7; GB:AE005172; NID:g4926828; PID:158168  
C:Genetics:  
A:Gene: T17H7.13  
A:Map position: 1

Query Match 55.6%; Score 5; DB 2; Length 415;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRH 8  
Db 65 WFLRH 69

RESULT 37  
I38027  
MLN 64 protein - human  
C:Species: Homo sapiens (man)  
C:Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 09-Jul-2004  
C:Accession: I38027; S60682  
R:Tomasetto, C.; Regnier, C.H.; Moog-Lutz, C.; Mattei, M.G.; Chenard, M.P.; Lidereau, R.;  
Genomics 28, 367-376, 1995  
A:Title: Identification of four novel human genes amplified and overexpressed in breast c  
A:Reference number: I37080; MUID:96039245; PMID:7490069  
A:Accession: I38027  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-445 <RES>  
A:Cross-references: UNIPROT:Q14849; UNIPARC:UPI000012F1BC; EMBL:X80198; NID:g951278; PID:158168  
A:Note: Submitted to the EMBL Data Library, July 1994  
C:Genetics:  
A:Gene: MLN64

Query Match 55.6%; Score 5; DB 2; Length 445;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

QY      2  ETWFL 6
      |||||
Db      165  ETWFL 169

RESULT 38
T40446
metaxin homolog - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T40446
R:Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.; Chillingworth, T.; Churcher, C.M.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21929
A:Accession: T40446
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-450 <L>N>
A:Cross-references: UNIPROT:Q9UUA5; UNIPARC:UPI0000069FAF; EMBL:AL109822; PIDN:CAB52621.
A:Experimental source: strain 972h-; cosmid c409
C:Genetics:
A:Gene: SPDB:SPBC409.17c
A:Map position: 2

      Query Match      55.6%; Score 5; DB 2; Length 450;
      Best Local Similarity 100.0%; Pred. No. 94;
      Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  RETWF 5
      |||||
Db      262  RETWF 266

RESULT 39
T06589
3-methyl-2-oxobutanoate dehydrogenase (lipoamide) (EC 1.2.4.4) E1-alpha chain precursor,
N:Alternate names: branched-chain alpha-keto acid dehydrogenase
C:Species: Lycopersicon esculentum (tomato)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06589
R:Giritch, A.; Baumlein, H.
submitted to the EMBL Data Library, April 1997
A:Description: A molecular cloning and characterization of cDNA coding for the branched
A:Reference number: Z15779
A:Accession: T06589
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-456 <GIR>
A:Cross-references: UNIPROT:O03849; UNIPARC:UPI00000A796F; EMBL:Z94180; PIDN:CAB08111.1
A:Experimental source: cultivar Bonner Beste, mutant chloronerva; roots
C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bin
C:Keywords: mitochondrion; oxidoreductase; phosphoprotein
F:1-34/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:35-456/Product: 3-methyl-2-oxobutanoate dehydrogenase (lipoamide) E1-alpha chain #stat
F:239-286/Domain: thiamin pyrophosphate-binding domain homology <TPB>

      Query Match      55.6%; Score 5; DB 2; Length 456;
      Best Local Similarity 100.0%; Pred. No. 95;
      Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3  TWFLR 7
      |||||
Db      3  TWFLR 7

RESULT 40
JC4313
keratin 16, type I, cytoskeletal - human
N:Alternate names: 46K keratin type I; cytokeratin 16
C:Species: Homo sapiens (man)
C>Date: 06-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C:Accession: JC4313; A24843; I58129
R:Paladini, R.D.; Takahashi, K.; Gant, T.M.; Coulombe, P.A.

```

```

Biochem. Biophys. Res. Commun. 215, 517-523, 1995
A:Title: cDNA cloning and bacterial expression of the human type I keratin 16.
A:Reference number: JC4313; MUID:96011809; PMID:7487986
A:Accession: JC4313
A:Molecule type: mRNA
A:Residues: 1-473 <PAL>
A:Cross-references: UNIPROT:P08779; UNIPARC:UPI000016B4A5; GB:S79867; NID:g1195530; PIDN:
A:Experimental source: epidermal keratinocytes
R:RayChaudhury, A.; Marchuk, D.; Lindhurst, M.; Fuchs, E.
Mol. Cell. Biol. 6, 539-548, 1986
A:Title: Three tightly linked genes encoding human type I keratins: conservation of sequ
A:Reference number: A24843; MUID:87064338; PMID:2431270
A:Accession: A24843
A:Molecule type: DNA
A:Residues: 1, 'T', 3-25, 'A', 27-37, 'A', 39-40, 43, 'ASTY', 48-49, 'A', 51-186, 'HAL', 190-207, 'ARTC
A:Cross-references: UNIPARC:UPI00001774B6; GB:M28433; NID:g186683; PIDN:AAAS9460.1; PID:
R:McLean, W.H.I.; Rugg, E.L.; Lunny, D.P.; Morley, S.M.; Lane, E.B.; Swensson, O.; Doppir
Kunkeler, L.; Munro, C.S.
Nature Genet. 9, 273-278, 1995
A:Title: Keratin 16 and keratin 17 mutations cause pachyonychia congenita.
A:Reference number: I58129; MUID:95291318; PMID:7539673
A:Accession: I58129
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 118-131, 'P', 133-134 <MCL>
A:Cross-references: UNIPARC:UPI000016B48F; GB:S78514; NID:g1000376; PIDN:AAB34564.1; PID:
A:Note: this is a mutant sequence
C:Comment: This protein is an intermediate filament protein and expressed in epithelial t
ound healing, psoriasis and cancer.
C:Genetics:
A:Gene: GDB:KRT16
A:Cross-references: GDB:136207; OMIM:148067
A:Map position: 17pter-17qter
A:Note: defects in this gene may result in Jadassohn-Lewandowsky pachyonychia congenita
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil

      Query Match      55.6%; Score 5; DB 2; Length 473;
      Best Local Similarity 100.0%; Pred. No. 98;
      Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  ETWFL 6
      |||||
Db      305  ETWFL 309

Search completed: August 31, 2006, 10:47:57
Job time : 24.25 secs

```



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:29:54 ; Search time 139.25 Seconds  
(without alignments)  
59.786 Million cell updates/sec

Title: DENGUE\_SEROTYPE3

Perfect score: 9

Sequence: 1 retwflrhp 9

Scoring table:  
OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2849398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database :

Uniprot 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	66.7	286	Q5FOA0	GLUOX
2	6	66.7	405	Q4IRY2	GIBBEZ
3	6	66.7	1079	Q1113	HUMAN
4	6	66.7	1079	Q53TA5	homo sapien
5	6	66.7	1093	Q8X0R0	NEUCR
6	6	66.7	1117	Q7S795	NEUCR
7	5	55.6	29	Q6CBE3	YARLW
8	5	55.6	77	Q8A2T7	BACTN
9	5	55.6	79	Q71136	LACDL
10	5	55.6	80	Q46AP0	METBA
11	5	55.6	80	Q8TIM1	METAC
12	5	55.6	87	Q5QCQ7	CENAS
13	5	55.6	95	Q9MIS5	STELE
14	5	55.6	101	Q3B3X5	PELID
15	5	55.6	108	Q37HR3	RHOPA
16	5	55.6	108	Q2J2M1	RHOPA
17	5	55.6	109	Q82LF3	STRAW
18	5	55.6	115	Y115	ADE02
19	5	55.6	115	Q2K522	ADE05
20	5	55.6	118	Q3WDX0	9ACTO
21	5	55.6	118	Q4QKE8	HAET8
22	5	55.6	122	Q5P5F8	AZOSE
23	5	55.6	124	Q47860	THEFY
24	5	55.6	127	1	CRCB_ERWCT
25	5	55.6	128	Q90Z26	XENTR
26	5	55.6	132	Q6TIW1	ANETH
27	5	55.6	132	Q4RTY3	TETNG
28	5	55.6	134	Q411L0	KINPA
29	5	55.6	136	Q977K8	9CREN
30	5	55.6	136	Q82HV8	STRAW
31	5	55.6	138	Q2WJ35	CLOBE
				Q2WJ35	clostridium

32	5	55.6	139	2	Q61GY3	DROME
33	5	55.6	139	2	Q4NBA0	9MICC
34	5	55.6	140	2	Q6ZOF6	ORYSA
35	5	55.6	141	2	Q67MA3	SYNTH
36	5	55.6	145	2	Q9RIK1	STRPY
37	5	55.6	157	2	Q849K6	STRVN
38	5	55.6	158	2	Q2PS61	9BACT
39	5	55.6	158	2	Q2PSC6	9BACT
40	5	55.6	158	2	Q2PSG0	9BACT
41	5	55.6	158	2	Q5KYB0	GEOKA
42	5	55.6	163	2	Q7NJ45	GLOVI
43	5	55.6	164	2	Q3VE10	9SPHN
44	5	55.6	166	2	Q2KC79	RHET
45	5	55.6	168	2	Q16641	CAEEL
46	5	55.6	170	2	Q34488	BACSU
47	5	55.6	171	2	Q2S721	9GAMW
48	5	55.6	171	2	Q8ZJZ4	SALTY
49	5	55.6	171	2	Q9EWQ7	STRCO
50	5	55.6	175	2	Q981I9	RHILLO
51	5	55.6	176	2	Q4MIV5	BACEE
52	5	55.6	176	2	Q737H4	BACC1
53	5	55.6	178	1	VNCA	RSVM
54	5	55.6	178	1	VNCA	RSVT
55	5	55.6	178	2	Q4TUA0	9VIRU
56	5	55.6	178	2	Q52P74	rice stripe
57	5	55.6	178	2	Q52R45	9VIRU
58	5	55.6	178	2	Q52R48	9VIRU
59	5	55.6	178	2	Q52R49	9VIRU
60	5	55.6	178	2	Q6EWP1	9VIRU
61	5	55.6	178	2	Q705B3	9VIRU
62	5	55.6	178	2	Q705B5	9VIRU
63	5	55.6	178	2	Q705B6	9VIRU
64	5	55.6	178	2	Q71TU0	9VIRU
65	5	55.6	178	2	Q71TL1	9VIRU
66	5	55.6	178	2	Q71TL2	9VIRU
67	5	55.6	178	2	Q71TL3	9VIRU
68	5	55.6	178	2	Q71TL5	9VIRU
69	5	55.6	178	2	Q71TL6	9VIRU
70	5	55.6	178	2	Q71TL7	9VIRU
71	5	55.6	178	2	Q71TL8	9VIRU
72	5	55.6	178	2	Q71TL9	9VIRU
73	5	55.6	178	2	Q80A47	9VIRU
74	5	55.6	178	2	Q80A50	9VIRU
75	5	55.6	178	2	Q91CC6	9VIRU
76	5	55.6	178	2	Q9JOW5	9VIRU
77	5	55.6	178	2	Q9JOW6	9VIRU
78	5	55.6	178	2	Q9JOW7	9VIRU
79	5	55.6	178	2	Q5K015	9VIRU
80	5	55.6	178	2	Q10389	9VIRU
81	5	55.6	178	2	Q5K003	9VIRU
82	5	55.6	178	2	Q5K011	9VIRU
83	5	55.6	178	2	Q5K013	9VIRU
84	5	55.6	183	2	Q5K013	9VIRU
85	5	55.6	190	2	Q6L1R8	PACTO
86	5	55.6	190	2	Q2UK03	ASPOR
87	5	55.6	191	2	Q3QVT3	9RHO
88	5	55.6	194	2	Q98B61	RHILLO
89	5	55.6	196	2	Q98R28	MYCPU
90	5	55.6	198	2	Q3GC97	9FIRM
91	5	55.6	198	2	Q82P58	SALTY
92	5	55.6	204	2	Q3F261	9BURK
93	5	55.6	204	2	Q3G826	9DEL
94	5	55.6	207	2	Q5NCE1	MOUSE
95	5	55.6	211	2	Q8R221	MOUSE
96	5	55.6	213	2	Q48B30	PSE14
97	5	55.6	213	2	Q6VEB2	PSE5Y
98	5	55.6	215	2	Q5FJB1	LACAC
99	5	55.6	218	2	Q5Z317	NOCAF
100	5	55.6	219	2	Q47QAI	THEFY
101	5	55.6	220	2	Q3W859	9ACTO
102	5	55.6	220	2	Q4OVL6	KINRA
103	5	55.6	221	2	Q9X802	STRCO
104	5	55.6	224	2	Q6F256	MESFL

Q61GY3	drosophila
Q4NBA0	arthrobacte
Q6ZOF6	oryza sativ
Q67MA3	symbiobacte
Q9RIK1	streptococc
Q849K6	streptococc
Q2PS61	uncultured
Q2PSC6	uncultured
Q2PSG0	uncultured
Q5KYB0	geobacillus
Q7NJ45	gloeobacter
Q3VE10	spingopyxi
Q2KC79	rhizobium e
Q16641	caenorhabdi
Q34488	bacillus su
Q2S721	hahella che
Q8ZJZ4	salmonella
Q9EWQ7	streptomyce
Q981I9	rhizobium l
Q4MIV5	bacillus ce
Q737H4	rice stripe
Q01209	rice stripe
Q00844	rice stripe
Q4TUA0	rice stripe
Q52P74	rice stripe
Q52R45	rice stripe
Q52R48	rice stripe
Q52R49	rice stripe
Q6EWP1	rice stripe
Q705B3	rice stripe
Q705B5	rice stripe
Q705B6	rice stripe
Q71TU0	rice stripe
Q71TL1	rice stripe
Q71TL2	rice stripe
Q71TL3	rice stripe
Q71TL5	rice stripe
Q71TL6	rice stripe
Q71TL7	rice stripe
Q71TL8	rice stripe
Q71TL9	rice stripe
Q80A47	rice stripe
Q80A50	rice stripe
Q91CC6	rice stripe
Q9JOW5	rice stripe
Q9JOW6	rice stripe
Q9JOW7	rice stripe
Q5K015	rice stripe
Q10389	rice stripe
Q5K003	rice stripe
Q5K011	rice stripe
Q5K013	rice stripe
Q5K013	rice stripe
Q6L1R8	picrophilus
Q2UK03	aspergillus
Q3QVT3	silicibacte
Q98B61	rhizobium l
Q98R28	mycoplasma
Q3GC97	synrophomo
Q82P58	salmonella
Q3F261	burkholderi
Q3G826	pelobacter
Q5NCE1	mus musculu
Q8R221	mus musculu
Q48B30	pseudomonas
Q6VEB2	pseudomonas
Q5FJB1	lactobacilli
Q5Z317	nocardia fa
Q47QAI	thermobifid
Q3W859	frankia sp.
Q4OVL6	kineococcus
Q9X802	streptomyce
Q6F256	mesoplasma

```

105 5 55.6 224 2 Q9A3B2 CAUCR
106 5 55.6 225 2 Q31V9 METHUS
107 5 55.6 225 2 Q4H823_9DSIO
108 5 55.6 231 2 Q4BWG1 CROWT
109 5 55.6 232 2 Q85SH1 ORYZA
110 5 55.6 232 2 Q2T527 BURTH
111 5 55.6 234 1 MENTO HUMAN
112 5 55.6 235 1 MENTO MOUSE
113 5 55.6 235 2 Q57NW5 SALCH
114 5 55.6 235 2 Q5PIM7 SALPA
115 5 55.6 235 2 Q827C5 SALTI
116 5 55.6 235 2 Q3U8Q7 MOUSE
117 5 55.6 235 2 Q3U8Q7 MOUSE
118 5 55.6 235 2 Q5U205 RAT
119 5 55.6 237 2 Q44AD1 SOLUS
120 5 55.6 238 2 Q7ND55 GLOVI
121 5 55.6 239 2 Q8TVN1 METHKA
122 5 55.6 244 2 Q16195 HUMAN
123 5 55.6 244 2 Q34170_9RHIZ
124 5 55.6 244 2 Q7D183 AGROBATERI
125 5 55.6 245 2 Q93SE2 ECOLI
126 5 55.6 245 2 Q8XX11 SYNPD
127 5 55.6 245 2 Q82GS8 STRAW
128 5 55.6 247 1 DDPX SYN3
129 5 55.6 248 2 Q5M912 XENTR
130 5 55.6 249 2 Q82F47 STRAM
131 5 55.6 251 2 Q6MSK2 MYCMS
132 5 55.6 253 2 Q53T67 HUMAN
133 5 55.6 254 2 Q5D6B8 PRRSV
134 5 55.6 256 2 Q9L0W3 STRCO
135 5 55.6 257 1 Y130 BUCBP
136 5 55.6 259 2 Q31W10 RHOS4
137 5 55.6 262 2 Q474R5 RALEJ
138 5 55.6 263 2 Q5DB61 SCHJA
139 5 55.6 265 2 Q8S5C9 ORYSA
140 5 55.6 270 1 YDHT ECOLI
141 5 55.6 270 2 Q31C6 SHIBS
142 5 55.6 270 2 Q32FA2 SHIDS
143 5 55.6 270 2 Q32FA2 SHIDS
144 5 55.6 270 2 Q32222 SHISS
145 5 55.6 270 2 Q8X618 ECOL6
146 5 55.6 270 2 Q8FH64 ECOL6
147 5 55.6 270 2 Q83KX2 SHIFL
148 5 55.6 273 2 Q4T0Z0 TETNG
149 5 55.6 273 2 Q4TF89 TETNG
150 5 55.6 276 2 Q3QME4_9GAMM

```

## ALIGNMENTS

```

RESULT 1
Q5FQA0 GLUOX PRELIMINARY; PRT; 266 AA.
ID Q5FQA0 GLUOX integrated into UniProtKB/TrEMBL.
AC Q5FQA0;
DT 01-MAR-2005, sequence version 1.
DE Putative hydrolase of the HAD superfamily.
EN OrderedLocusNames=COX1706;
OS Gluconobacter oxydans (Gluconobacter suboxydans).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconobacter.
OX NCBI_TaxID=442;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=621H;
RX PubMed=15665824; DOI=10.1038/nbr1062;
RA Prust C., Hoffmeister M., Liesegang H., Wiezer A., Fricke W.P.,
RA Ehrenreich A., Gottschalk G., Deppenmeier U.;
RT "Complete genome sequence of the acetic acid bacterium Gluconobacter
RT oxydans."
RL Nat. Biotechnol. 23:195-200(2005).

```

```

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000009; AAW61446.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR013200; HAD_3
DR InterPro; IPR006379; HAD_SF_IIB.
DR InterPro; IPR000150; Hypothet_cof.
DR Pfam; PF00702; Hydrolase; 1.
DR TIGRFAMs; TIGR00099; Cof-subfamily; 1.
DR TIGRFAMs; TIGR01484; HAD-SF-IIB; 1.
DR PROSITE; PS01229; COF_2; UNKNOWN_1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 266 AA; 28559 MW; A20E08223C537EFE CRC64;

Query Match 66.7%; Score 6; DB 2; Length 266;
Best Local Similarity 100.0%; Pred.No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETWFLR 7
Db 98 ETWFLR 103

RESULT 2
Q4IRY2 GIBZE PRELIMINARY; PRT; 405 AA.
ID Q4IRY2;
AC Q4IRY2;
DT 16-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN ORFNames=FG00026.1;
OS Gibberella zeae (Fusarium graminearum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=5518;
[1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
STRAIN=PH-1 / NRRL 31084;
RA Birren N., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepl Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D.,
RA Galagan J.E., Gardyna S., Gnerre S., Graham L., Grand-Pierre N.,
RA Hafez N., Hagopian D., Hagos B., Hall J., Horton L., Hulme W.,
RA Iliev I., Jaffe D., Johnson R., Jones C., Kamal M., Kamat A.,
RA Karatas A., Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G.,
RA Lui A., Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J.,
RA Manning J., Matthews C., Mauceli E., McCarthy M., Meldrim J.,
RA Meneus L., Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C.,
RA Nicol R., Nielsen C.B., Norbu C., O'Connor T., O'Donnell P.,
RA O'Neill D., Oliver J., Peterson K., Phunkhang P., Pierre N.,
RA Purcell S., Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C.,
RA Rogov P., Roman J., Schauer S., Schupbach R., Seaman S., Severy P.,
RA Smirnov S., Smith C., Spencer B., Stange-Thomann N., Stojanovic N.,
RA Stubbs M., Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.S.;
RT "Fusarium graminearum genome sequence."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

```



```

FT HTGWMARKTGLPEK (in isoform 2).
FT /FTId=VSP_012815.
FT Missing (in isoform 2).
FT /FTId=VSP_012815.
FT VSCCLQILSCASKSMEGIPWPSBDMGTARS -> ATNEG
FT CILEHSGGSDTARKTDASE (in isoform 2).
FT /FTId=VSP_012817.
FT A -> T (in dbSNP:2052937).
FT /FTId=VAR_024475.
FT A -> V (in Ref. 3; AAQ88539).
FT H -> Y (in Ref. 3; AAQ88539).
FT M -> T (in Ref. 3; AAQ88581).
FT M -> T (in Ref. 3; AAQ88581).
SQ SEQUENCE 1079 AA; 116341 MW; A18CA158F4DDBB9C CRC64;

Query Match 66.7%; Score 6; DB 1; Length 1079;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9
Db 267 WFLRHP 272

RESULT 4
Q53TA5 HUMAN PRELIMINARY; PRT; 1079 AA.
AC Q53TA5_5, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein GPR113.
GN Name=GPR113;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Swearingen S., Cordes M., Cotton M.;
RT "The sequence of Homo sapiens BAC clone RP11-499P9.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AC010896; AAY14645.1; -; Genomic DNA.
DR Ensembl; ENSG00000173567; Homo sapiens.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR GO; GO:0007218; P:neuropeptide signaling pathway; IEA.
DR InterPro; IPR013032; EGF_like_reg.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormone_rcpt.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS_1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR SMART; SM00303; GPS; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN 1.
DR PROSITE; PS00650; G_PROTEIN_RECFP_F2_2; UNKNOWN_1.

```

```

DR PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS50221; GFS; 1.
KW Hypothetical protein.
SQ SEQUENCE 1079 AA; 116341 MW; A18CA158F4DDBB9C CRC64;

Query Match 66.7%; Score 6; DB 2; Length 1079;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9
Db 267 WFLRHP 272

RESULT 5
Q8XOR0 NEUCR PRELIMINARY; PRT; 1093 AA.
AC Q8XOR0;
DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2002, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Hypothetical protein SE6.080.
GN Name=SE6.080;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Schulte U., Align V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA German Neurospora genome project;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AL670004; CAD21248.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 1093 AA; 120695 MW; 9F6BF07A8AD661BD CRC64;

Query Match 66.7%; Score 6; DB 2; Length 1093;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9
Db 621 WFLRHP 626

RESULT 6
Q7S795 NEUCR PRELIMINARY; PRT; 1117 AA.
AC Q7S795;
DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT 15-DEC-2003, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Predicted protein.
GN ORFNames=NCU08869.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=22598136; PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,

```

RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Ianakiev P., Bell-Pedersen D., Nelson M.A.,  
 RA Werner-Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L.,  
 RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,  
 RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,  
 RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamvysseilis M.,  
 RA Mauceli E., Bielek C., Rued S., Frishman D., Kryzstofova S.,  
 RA Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S., Cogoni C.,  
 RA Macino G., Catchside D.E.A., Li W., Pratt R.J., Osmani S.A.,  
 RA DeSouza C.P.C., Glass N.L., Orbach M.J., Berglund J.A., Voelker R.,  
 RA Yarden O., Plamann M., Seiler S., Dunlap J.C., Radford-A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannheim J., Ebbola D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.W.,  
 RA "The genome sequence of the filamentous fungus *Neurospora crassa*."  
 RL Nature 422:859-868(2003).  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL: AABX01000300; EAA31426.1; -; Genomic DNA  
 DR SEQUENCE 1117 AA; 123343 MW; 198B80EC607752D CRC64;  
 SQ  
 Query Match 66.7%; Score 6; DB 2; Length 1117;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 WFLRHP 9  
 DB 621 WFLRHP 626  
 RESULT 7  
 ID Q6CGE3\_YARLI PRELIMINARY; PRT; 29 AA.  
 AC Q6CGE3;  
 DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.  
 DT 16-AUG-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 12.  
 DE Similarity.  
 GN OrderedLocNames=VAL10A20042g;  
 OS Yarrowia lipolytica (Candida lipolytica).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Dipodascaceae; Yarrowia.  
 OX NCBI\_TaxID=4952;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=CLIB 122 / E 150;  
 RX PubMed=15229592; DOI=10.1038/nature02579;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barney S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,  
 RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.-M., Nikolski M., Oztas S., Olier-Kalogoropoulos O.,  
 RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,  
 RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.-L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL: CR382127; CAG84207.1; -; Genomic DNA.  
 DR Complete proteome.  
 KW

SQ SEQUENCE 29 AA; 3169 MW; 5FB2DCF7AA4626ED CRC64;  
 Query Match 55.6%; Score 5; DB 2; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRHP 9  
 DB 15 FLRHP 19  
 RESULT 8  
 ID Q8A2T7\_BACTN PRELIMINARY; PRT; 77 AA.  
 AC Q8A2T7;  
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2003, sequence version 1.  
 DT 07-FEB-2006, entry version 9.  
 DE Hypothetical protein.  
 GN OrderedLocNames=BT3218; ORFNames=BT\_3218;  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;  
 OC Bacteroidaceae; Bacteroides.  
 OX NCBI\_TaxID=818;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=VPI-5482 / ATCC 29148;  
 RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;  
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
 RA Chiang H.C., Hooper L.V., Gordon J.I.;  
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."  
 RL Science 299:2074-2076(2003).  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL: AB015928; AAO78324.1; -; Genomic DNA.  
 DR Biocyc; BTHE226186; BT3218-MONOMER; -;  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 77 AA; 9164 MW; 115052AB1896BA18 CRC64;  
 Query Match 55.6%; Score 5; DB 2; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRHP 9  
 DB 66 FLRHP 70  
 RESULT 9  
 ID Q7I136\_LACDL PRELIMINARY; PRT; 79 AA.  
 AC Q7I136;  
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.  
 DT 05-JUL-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 7.  
 DE Sufi protein (Fragment)  
 OS Lactobacillus delbrueckii subsp. lactis.  
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
 OC Lactobacillus.  
 OX NCBI\_TaxID=29397;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ATCC 47997;  
 RA Langenheim J.F., Ulrich R.L.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL: AF496414; AAO07102.1; -; Genomic DNA.  
 DR NON\_TER 1  
 FT 1

```

FT NON TER 79 79
SQ SEQUENCE 79 AA; 9042 MW; 5B609E1F96CFDFD10 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 44 FLRHP 48

RESULT 10
Q46AP0.METBA PRELIMINARY; PRT; 80 AA.
AC Q46AP0;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN OrderedLocusNames=Mbar_A2121;
OS Methanosarcina barkeri.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2208;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Fusaro / DSM 804;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hammon N., Istrani S., Pittluck S., Goodwin L.A., Saunders E.H.,
RA Schmutz J., Larimer F., Land M., Anderson I., Richardson P.;
RT "Complete sequence of chromosome 1 of Methanosarcina barkeri str.
RT Fusaro.";
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; CP000099; AAZ71052.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 80 AA; 9187 MW; 2DD0ED9A5E5C4CB3 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 32 FLRHP 36

RESULT 11
Q8TIM1.METAC PRELIMINARY; PRT; 80 AA.
AC Q8TIM1;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-MAR-2006, entry version 11.
DE Hypothetical protein.
GN ORFNames=MA_4126;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Gargan J.E., Nusbbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Achnor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,

```

```

RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett W.M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.M., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AE010299; AAM07474.1; -; Genomic_DNA.
DR GenomeReviews; AE010299 GR; MA4126.
DR Biocyc; MACEI88937:MA4126-MONOMER; -.
DR InterPro; IPR012933; Ycfa.
DR Pfam; PF07927; Ycfa; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 80 AA; 9184 MW; DF02721324F5D173 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 32 FLRHP 36

RESULT 12
Q5QCQ7.CENAS PRELIMINARY; PRT; 87 AA.
ID Q5QCQ7.CENAS
AC Q5QCQ7;
DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein (fragment).
OS Centibacterium arsenoxidans.
OC Bacteria; Centibacterium.
OX NCBI_TaxID=204773;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=ULPAS1.
RA Carapito C., Muller D., Turlin E., Riegel P., Leize E., Danchin A.,
RA Van Dorsselaer A., Bertin P., Lett M.-C.,
RT "Pleiotropic effect of arsenic stress on Centibacterium arsenoxidans, a
RT metalloresistant beta-proteobacterium.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AY728027; AAV68356.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON TER 87
SQ SEQUENCE 87 AA; 9653 MW; EAAD40B00A2E3C86 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8
Db 20 WFLRH 24

RESULT 13
Q9MIS5.9TELE PRELIMINARY; PRT; 95 AA.
ID Q9MIS5.9TELE
AC Q9MIS5;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.

```

DT 07-FEB-2006, entry version 21.  
DE Cytochrome b (Fragment).  
GN Name=cytb;  
OS Retropinna tasmanica.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Retropinnidae; Retropinna.  
OX NCBI\_TaxID=89573;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22111806; PubMed=12116439; DOI=10.1080/106351500750049824;  
RA Waters J.M., Lopez J.A., Wallis G.P.;  
RT "Molecular phylogenetics and biogeography of galaxiid fishes  
RT (Osteichthyes: Galaxiidae): dispersal, vicariance and the position of  
RT Lepidogalaxias salamandroides.";  
RL Syst. Biol. 49:777-795(2000).  
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase  
CC complex (complex III or cytochrome b-c1 complex), which is a  
CC respiratory chain that generates an electrochemical potential  
CC coupled to ATP synthesis (By similarity).  
CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).  
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,  
CC cytochrome c1 and the Rieske protein (By similarity).  
CC -1- SIMILARITY: Belongs to the cytochrome b family.  
CC  
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
CC EMBL: AF112321; AAF67414.1; -; Genomic\_DNA.  
DR SMR; Q9MIS5; 1-95.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0016020; C: membrane; IEA.  
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.  
DR GO; GO:0005739; C: mitochondrion; IEA.  
DR GO; GO:0005506; F: iron ion binding; IEA.  
DR GO; GO:0046872; F: metal ion binding; IEA.  
DR GO; GO:0016491; F: oxidoreductase activity; IEA.  
DR GO; GO:0006118; P: electron transport; IEA.  
DR InterPro; IPR005797; Cytb\_b6\_N.  
DR Pfam; PF00033; Cytochrom\_b\_N\_T\_1.  
DR PROSITE; PS1002; CYTB\_NTER; 1.  
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;  
KW Respiratory chain; Transmembrane; Transport.  
FT NON\_TER 1 95  
FT TER 95  
SQ SEQUENCE 95 AA; 10578 MW; E7F5ABDD28E269DE CRC64;  
  
Query Match 55.6%; Score 5; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 ETWFL 6  
DB 77 ETWFL 81  
  
RESULT 14  
Q3BX5\_PELLD PRELIMINARY; PRT; 101 AA.  
ID Q3BX5\_PELLD  
AC Q3BX5;  
DT 22-NOV-2005, integrated into UniprotKB/TREMBL.  
DT 22-NOV-2005, sequence version 1.  
DT 07-FEB-2006, entry version 3.  
DE Hypothetical protein.  
GN ORFNames=Plut\_1094;  
OS Pelodictyon luteolum (strain DSM 273) (Chlorobium luteolum (strain DSM  
OS 273)).  
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
OC Chlorobium/Pelodictyon group; Pelodictyon.  
OX NCBI\_TaxID=319225;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=DSM 273;  
RG US DOE Joint Genome Institute;  
RA Copeland A., Lucas S., Lapidus A., Barty K., Detter J.C., Glavina T.,  
RA Hammon N., Israni S., Pitluck S., Bryant D., Schmutz J., Larimer F.,  
RA Land M., Kyripides N., Ivanova N., Richardson P.;  
RT "Complete sequence of Pelodictyon luteolum DSM 273.";  
RL Submitted (AUG-2005) to the EMBL/GenBank/DBJ databases.  
CC  
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
CC EMBL: CP000096; AB23956.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 101 AA; 11173 MW; AAEF2D3DE11B891C CRC64;  
  
Query Match 55.6%; Score 5; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 TWFLR 7  
DB 26 TWFLR 30  
  
RESULT 15  
Q3HR3\_RHOHA PRELIMINARY; PRT; 108 AA.  
ID Q3HR3\_RHOHA  
AC Q3HR3;  
DT 06-DEC-2005, integrated into UniprotKB/TREMBL.  
DT 06-DEC-2005, sequence version 1.  
DT 07-FEB-2006, entry version 3.  
DE Hypothetical protein.  
GN ORFNames=RPDDRAFT\_1705;  
OS Rhodopseudomonas palustris BisB5.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Rhodopseudomonas.  
OX NCBI\_TaxID=316057;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=BisB5;  
RG US DOE Joint Genome Institute (JGI-PGF);  
RA Copeland A., Lucas S., Lapidus A., Barty K., Detter J.C., Glavina T.,  
RA Hammon N., Israni S., Pitluck S., Richardson P.;  
RT "Sequencing of the draft genome and assembly of Rhodopseudomonas  
RT palustris BisB5.";  
RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=BisB5;  
RG US DOE Joint Genome Institute (JGI-ORNL);  
RA Larimer F., Land M.;  
RT "Annotation of the draft genome of Rhodopseudomonas palustris BisB5.";  
RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
CC  
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC  
CC EMBL: AAKZ01000007; EAO85463.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 108 AA; 11908 MW; 5DD8C5B4A85507BD CRC64;  
  
Query Match 55.6%; Score 5; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RETWF 5  
DB 100 RETWF 104



```
RESULT 16
Q2J2M1 RHOPA PRELIMINARY; PRT; 108 AA.
AC Q2J2M1;
DT 07-MAR-2006, integrated into UniProtKB/TrEMBL.
DT 07-MAR-2006, sequence version 1.
DT 07-MAR-2006, entry version 1.
DE Hypothetical protein.
GN ORFNames=RPB_0578;
OS Rhodopseudomonas palustris Haa2.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=316058;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RP STRAIN=Haa2;
RG US DOE Joint Genome Institute;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hamon N., Israni S., Pitluck S., Chain P., Malfatti S., Shin M.,
RA Verges L., Schmutz J., Larimer F., Land M., Hauser L., Pelletier D.A.,
RA Kyrpides N., Anderson I., Oda Y., Harwood C.S., Richardson P.;
RT "Complete sequence of Rhodopseudomonas palustris Haa2.";
RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000250; ABD05289.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 12049 MW; 235632516D148B20 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETWF 5
Db 100 RETWF 104

RESULT 17
Q82LF3 STRAW PRELIMINARY; PRT; 109 AA.
AC Q82LF3;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Hypothetical protein.
GN OrderedLocusNames=SAV2057;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary
metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
```

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BA000030; BAC69768.1; -; Genomic_DNA.
DR Biocyc; SAV227882:SAV2057-MONOMER; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
KW Complete proteome; DNA-binding; Hypothetical protein; Transcription;
KW Transcription regulation.
SQ SEQUENCE 109 AA; 12127 MW; 9BF1F50C411DFAD2 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 77 FLRHP 81

RESULT 18
Y115 ADE02 STANDARD; PRT; 115 AA.
ID Y115 ADE02
AC P03290;
DT 21-JUL-1986, integrated into UniProtKB/Swiss-Prot.
DT 21-JUL-1986, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Hypothetical protein E-115.
OS Human adenovirus 2 (HAdV-2).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10515;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=83056843; PubMed=7142161;
RA Gingeras T.R., Sciaky D., Gellinas R.E., Bing-Dong J., Yen C.E.,
RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;
RT "Nucleotide sequences from the adenovirus-2 genome.";
RL J. Biol. Chem. 257:13475-13491(1982).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; J01917; -; NOT ANNOTATED_CDS; Genomic_DNA.
DR PIR; A03862; A03862.
KW Hypothetical protein.
FT CHAIN 1 115 Hypothetical protein E-115.
FT SEQUENCE 115 AA; 12236 MW; C7A08EA239B8FD98 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ETWFL 6
Db 3 ETWFL 7

RESULT 19
Q2KS22 ADE05 PRELIMINARY; PRT; 115 AA.
ID Q2KS22 ADE05
AC Q2KS22;
DT 07-MAR-2006, integrated into UniProtKB/TrEMBL.
DT 07-MAR-2006, sequence version 1.
DT 07-MAR-2006, entry version 1.
DE Hypothetical 12 kDa early protein.
OS Human adenovirus 5 (HAdV-5).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28285;
RN [1]
```



RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN-NHRC AdSPS 7151;  
 RG Epidemic Outbreak Surveillance (EOS);  
 RA Tibbets C., Purkayastha A., Su J., Russell K., Carlisle S.,  
 RA Opina R., Reynolds T., Rowley R., Hanson E., Seto D.,  
 RT "The complete nucleotide sequence and genome organization of Human  
 RT adenovirus serotype 5, field strain.";  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR ENBL: AY601635; AAW65500.1; -: Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 115 AA; 12210 MW; DF1B2DA239AA7F08 CRC64;  
 Query Match 55.6%; Score 5; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ETWFL 6  
 Db 3 ETWFL 7  
 RESULT 20  
 Q3WDX0\_9ACTO PRELIMINARY; PRT; 118 AA.  
 ID Q3WDX0\_9ACTO PRELIMINARY; PRT; 118 AA.  
 AC Q3WDX0;  
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
 DT 11-OCT-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 3.  
 DE Putative cytochrome P450.  
 GN ORFNames=FraneanIDRAFT\_5182;  
 OS Frankia sp. EAN1pec.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Frankineae; Frankiaceae; Frankia.  
 OC NCBI\_TaxID=298653;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=EAN1pec;  
 RG US DOE Joint Genome Institute (JGI-PGF);  
 RA Copeland A., Lucas S., Lepidus A., Barry K., Dettler C., Glavina T.,  
 RA Hammon N., Israni S., Pitluck S., Richardson P.;  
 RA "Sequencing of the draft genome and assembly of Frankia sp. EAN1pec.";  
 RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=EAN1pec;  
 RG US DOE Joint Genome Institute (JGI-ORNL);  
 RA Larimer F., Land M.;  
 RT "Annotation of the draft genome assembly of Frankia sp. EAN1pec.";  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR ENBL: AAI101000014; EAN16946.1; -: Genomic\_DNA.  
 DR GO: GO:0020037; F:heme binding; IEA.  
 DR GO: GO:0005506; F:iron ion binding; IEA.  
 DR GO: GO:0004497; F:monooxygenase activity; IEA.  
 DR GO: GO:0006118; P:electron transport; IEA.  
 DR InterPro: IPR002397; BP450.  
 DR InterPro: IPR001128; Cytochrome\_P450.  
 DR PRINTS: PR00359; BP450.  
 SQ SEQUENCE 118 AA; 13148 MW; 8EEA8775EFB424AD CRC64;  
 Query Match 55.6%; Score 5; DB 2; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9  
 Db 61 FLRHP 65  
 RESULT 21  
 Q4QKE8\_HAE18 PRELIMINARY; PRT; 118 AA.  
 ID Q4QKE8\_HAE18 PRELIMINARY; PRT; 118 AA.  
 AC Q4QKE8;  
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.  
 DT 19-JUL-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 5.  
 DE Putative integrase/recombinase.  
 GN OrderedLocusNames=NT11711;  
 OS Haemophilus influenzae (strain 86-028NP).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OC NCBI\_TaxID=281310;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RP PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;  
 RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,  
 RA Zhong H., Gibson J., Gibson M., Johnson L.S., Lewis L., Bakalatz L.O.,  
 RA Munson R.S. Jr.;  
 RA "Genomic sequence of an otitis media isolate of nontypeable  
 RT Haemophilus influenzae: comparative study with H. influenzae serotype  
 RT d, strain KW20.";  
 RL J. Bacteriol. 187:4627-4636(2005).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR ENBL: CP000057; AAX88499.1; -: Genomic\_DNA.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0015074; P:DNA integration; IEA.  
 DR GO: GO:0006310; P:DNA recombination; IEA.  
 KW Complete proteome.  
 SQ SEQUENCE 118 AA; 13822 MW; 9FCB660420C82E38 CRC64;  
 Query Match 55.6%; Score 5; DB 2; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRHP 9  
 Db 21 FLRHP 25  
 RESULT 22  
 Q5P5F8\_AZOSE PRELIMINARY; PRT; 122 AA.  
 ID Q5P5F8\_AZOSE PRELIMINARY; PRT; 122 AA.  
 AC Q5P5F8;  
 DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.  
 DT 04-JAN-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 8.  
 DE Hypothetical protein.  
 GN OrderedLocusNames=AZOSEAL3290; ORFNames=eba2387;  
 OS Azorarcus sp. (strain EbN1).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;  
 OC Rhodocyclaceae; Azorarcus.  
 OC NCBI\_TaxID=76114;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RP PubMed=15551059; DOI=10.1007/s00203-004-0742-9;  
 RA Rabus R., Kube M., Heider J., Beck A., Heitmann K., Widdel F.,  
 RA Reinhardt R.;  
 RT "The genome sequence of an anaerobic aromatic-degrading denitrifying  
 RT bacterium, strain EbN1.";  
 RL Arch. Microbiol. 183:27-36(2005).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License

```
CC -----
DR EMBL; CR555306; CAI07454.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 122 AA; 13581 MW; 18B790A94ECD3255 CRC64;

Query Match
Best Local Similarity 55.6%; Score 5; DB 2; Length 122;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 16 FLRHP 20

RESULT 23
Q47S60 THEFY PRELIMINARY; PRT; 124 AA.
AC Q47S60;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN OrderedLocusNames=Tfu_0669;
OS Thermobifida fusca (strain YX).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptoporangineae; Nocardiopsaceae; Thermobifida.
OX NCBI_TaxID=269800;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG US DOE Joint Genome Institute;
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Di Bartolo G., Chain P., Schmutz J.,
RA Larmer F., Land M., Lykidis A., Richardson P.;
RT "Complete sequence of Thermobifida fusca YX."
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000088; AAZ54707.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 124 AA; 14456 MW; 5F749F9A86A83FC0 CRC64;

Query Match
Best Local Similarity 55.6%; Score 5; DB 2; Length 124;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRH 8
Db 115 WFLRH 119

RESULT 24
CRCB ERWCT STANDARD; PRT; 127 AA.
AC Q6DJN0;
DT 05-JUL-2005, integrated into UniProtKB/Swiss-Prot.
DT 16-AUG-2004, sequence version 1.
DT 07-MAR-2006, entry version 15.
DE Protein crCB homolog.
GN Names=crCB; OrderedLocusNames=ECA1295;
OS Erwinia carotovora subsp. atroseptica (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebaiha M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
```

```
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- SUBCELLULAR LOCATION: Bacterial cell inner membrane; multi-pass
CC membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the crCB family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BX950851; CAG74205.1; -; Genomic DNA.
DR GenomeReviews; BX950851_GR; ECA1295.
DR HAMAP; MF_00454; -; 1.
DR InterPro; IPR003691; Camphor_CrCB.
DR Pfam; PF02537; CRCB; 1.
DR TIGRFAMs; TIGR00494; crCB; 1.
KW Complete proteome; Inner membrane; Membrane; Transmembrane.
FT CHAIN 1 127 Protein crCB homolog.
FT TRANSMEM 4 24 Potential.
FT TRANSMEM 35 55 Potential.
FT TRANSMEM 71 91 Potential.
FT TRANSMEM 103 123 Potential.
SQ SEQUENCE 127 AA; 13391 MW; ADED63C701397633 CRC64;

Query Match
Best Local Similarity 55.6%; Score 5; DB 1; Length 127;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 55 FLRHP 59

RESULT 25
Q90Z26 XENTR PRELIMINARY; PRT; 128 AA.
AC Q90Z26;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Xcat-2.
GN Name=Xcat-2;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Vempati U.D., King M.L.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF256086; AAK49295.1; -; mRNA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006445; P:regulation of translation; IEA.
DR InterPro; IPR008705; Nanos_RNA_bd.
DR Pfam; PF05741; zf-nanos; 1.
SQ SEQUENCE 128 AA; 14140 MW; E79556DEF1C0880B CRC64;

Query Match
Best Local Similarity 55.6%; Score 5; DB 2; Length 128;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
Db 111
```

```

Db      117 FLRHP 121
RESULT 26
Q6TIW1 ANETH PRELIMINARY; PRT; 132 AA.
AC Q6TIW1-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 16.
DE Putative transposase.
OS Aneurinibacillus thermoaerophilus.
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae;
OC Aneurinibacillus group; Aneurinibacillus.
OX NCBI_TaxID=143495;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC SFRAIN=L420-91T;
RX PubMed=15044388; DOI=10.1093/glycob/cwh064;
RA Schaffer C., Messner P.;
RT "Surface-layer glycoproteins: an example for the diversity of
RT bacterial glycosylation with promising impacts on nanobiotechnology.";
RL Glycobiology 14:31R-42R(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L420-91T;
RX PubMed=15316277;
RA Novotny R., Pfoestl A., Messner P., Schaffer C.;
RT "Genetic organization of chromosomal S-layer glycan biosynthesis loci
RT of Bacillaceae.";
RL Glycoconj. J. 20:435-447(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR ENBL; AY442352; AAS55727.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR InterPro; IPR012337; RNaseH_fold.
KW Hydrolase; Nuclease.
SQ SEQUENCE 132 AA; 15708 MW; 603062293C9D57B0 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TWFLR 7
Db 25 TWFLR 29

RESULT 27
Q4RTY3 TETNG PRELIMINARY; PRT; 132 AA.
AC Q4RTY3-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome 12 SCAP14996, whole genome shotgun sequence.
GN ORFNames=GSTENG00029043001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,

```

```

RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Scallius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR ENBL; CAAE01014996; CAG08149.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0008092; F:cytoskeletal protein binding; IEA.
DR InterPro; IPR000299; Band_4.1.
DR Pfam; PF00373; Band_4.1.
DR PRINTS; PR00661; ERMFAMILY.
DR PROSITE; PS50057; FERM_3; 1.
SQ SEQUENCE 132 AA; 15482 MW; 523BE6E225D67CFC CRC64;

Query Match 55.6%; Score 5; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RETWF 5
Db 57 RETWF 61

RESULT 28
Q411L0 KINRA -PRELIMINARY; PRT; 134 AA.
AC Q411L0-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=KradRAFT_2276;
OS Kineococcus radiotolerans SRS30216.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Frankineae; Kineosporiaceae; Kineococcus.
OX NCBI_TaxID=266940;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SRS30216;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Piluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Kineococcus
RT radiotolerans SRS30216.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SRS30216;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of draft genome assembly of Kineococcus radiotolerans
RT SRS30216.";

```

```

RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SR30216;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: the sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAEF02000024; EAM74977.1; -; Genomic_DNA.
SQ SEQUENCE 134 AA; 15058 MW; F037BBE97A0D4676 CRC64;

Query Match          55.6%; Score 5; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 TWFLR 7
DB      41 TWFLR 45

RESULT 29
Q977K8 9CREN
ID Q977K8 9CREN PRELIMINARY; PRT; 136 AA.
AC Q977K8;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Secreted protein.
OS uncultured crenarchaeote 74A4.
OC Archaea; Crenarchaeota; environmental samples;
OC marine Archaeal group 1.
OX NCBI_TaxID=166279;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21633832; PubMed=11772643; DOI=10.1128/ASM.68.1.335-345.2002;
RA Beja O., Koonin E.V., Aravind L., Taylor L.T., Seitz H., Stein J.L.,
RA Bensen D.C., Feldman R.A., Swanson R.V., DeLong E.F.;
RT "Comparative Genomic Analysis of Archaeal Genotypic Variants in a
RT Single Population and in Two Different Oceanic Provinces.";
RL Appl. Environ. Microbiol. 68:335-345(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF393466; AAK96100.1; -; Genomic DNA.
SQ SEQUENCE 136 AA; 15922 MW; 852D6DD1B1626B5C CRC64;

Query Match          55.6%; Score 5; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 TWFLR 7
DB      61 TWFLR 65

RESULT 30
Q82HV8 STRAW
ID Q82HV8 STRAW PRELIMINARY; PRT; 136 AA.
AC Q82HV8;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Hypothetical protein.
GN OrderedLocusNames=SAV3400;

```

```

OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BA000030; BAC71112.1; -; Genomic_DNA.
DR BioCyc; SAVE227882; SAV3400-MONOMER; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 136 AA; 14797 MW; 8A1E1A1D59C1F6F3 CRC64;

Query Match          55.6%; Score 5; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 FLRHP 9
DB      11 FLRHP 15

RESULT 31
Q2WJ35 CLOBE
ID Q2WJ35_CLOBE PRELIMINARY; PRT; 138 AA.
AC Q2WJ35;
DT 10-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 10-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Ribonucleoside triphosphate reductase.
GN ORFNames=CheiDRAFT_0504.
OS Clostridium beijerincki NCIMB 8052.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=290402;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCIMB 8052;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Clostridium
RT beijerincki NCIMB 8052.";
RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCIMB 8052;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Clostridium
RT NCIMB 8052.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an

```

```

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAL001000040; EAF58092.1; -; Genomic DNA.
SQ SEQUENCE 138 AA; 16559 MW; 70DB16611E5307AC CRC64;

Query Match          55.6%; Score 5; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RETWF 5
Db 43 RETWF 47

RESULT 32
Q61GY3_DROME
ID Q61GY3_DROME PRELIMINARY; PRT; 139 AA.
AC Q61GY3;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE H0C04272.
GN ORFNames=HDC04272;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solov'yev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J.D.,
RA Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BK003633; DAA02331.1; -; Genomic DNA.
SQ SEQUENCE 139 AA; 14909 MW; DBA4F99D68E4045D CRC64;

Query Match          55.6%; Score 5; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9
Db 78 FLRHP 82

RESULT 33
Q4NBA0_9MICC
ID Q4NBA0_9MICC PRELIMINARY; PRT; 139 AA.
AC Q4NBA0;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-MAR-2006, entry version 5.
DE Similar to Glutaredoxin and related proteins.
GN ORFNames=ArthDRAFT_0289;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcales; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RN Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RL [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PGF-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAG01000023; EAL94631.1; -; Genomic DNA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR011915; GLRX acting.
DR InterPro; IPR012336; Thioridoxin-like fd.
DR InterPro; IPR006662; Thioridoxin.
DR InterPro; IPR006663; Thioridoxin_dom2.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMS; TIGR02200; GLRX actino; 1.
DR PROSITE; PS00194; THIOREDOXIN; UNKNOWN 1.
SQ SEQUENCE 139 AA; 14842 MW; 188B6C4668B8D6E6 CRC64;

Query Match          55.6%; Score 5; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRHP 9
Db 19 FLRHP 23

RESULT 34
Q620F6_ORYSA
ID Q620F6_ORYSA PRELIMINARY; PRT; 140 AA.
AC Q620F6;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein OSUNBa0062G05.18.
GN Name=OSUNBa0062G05.18;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade;
OC Ehrhartoideae; Oryzoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
RT clone:OSUNBa0062G05.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF005491; BAD03650.1; -; Genomic DNA.
DR Gramene; Q620F6; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003822; PAH.
DR Pfam; PF02671; PAH; 1.
DR Hypothetical protein.
RW

```

```

GO; CO:0005525; F:GTP binding; IEA.
DR InterPro; IPR006074; GTP1_OBG_dom.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR006169; GTP1_OBG_sub.
DR InterPro; IPR002917; MMR_HSR1_GTP_bd.
DR Pfam; PF01018; GTP1_OBG_1.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00326; GTP1_OBG.
DR PROSITE; PS00905; GTP1_OBG; 1.
FT NON_TER 1
FT TER 145
SQ SEQUENCE 145 AA; 15297 MW; 3FE9DC259C5A3986 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 145;
Best Local Similarity 100.0%; Pred.No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 140 FLRHP 144

RESULT 37
ID Q849K6_STRVN PRELIMINARY; PRT; 157 AA.
AC Q849K6;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein psv2.34c.
GN Names=psv2.34c;
OS Streptomyces violaceoruber.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1935;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SANK95570;
RX MEDLINE=22123362; PubMed=12127493;
RA Spatz K., Kohn H., Redenbach M.;
RT "Characterization of the Streptomyces violaceoruber SANK95570 plasmids
RT psv1 and psv2.";
RL FEMS Microbiol. Lett. 213:87-92(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SANK95570;
RA Spatz K., Scholz C.J., Redenbach M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
CC EMBL; AY211023; AA050118.1; -; Genomic_DNA.
DR DR
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 157 AA; 17330 MW; AF4439D5B1270D24 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 157;
Best Local Similarity 100.0%; Pred.No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRHP 9
DB 96 FLRHP 100

RESULT 38
Q2PS61_9BACT
ID Q2PS61_9BACT PRELIMINARY; PRT; 158 AA.
AC Q2PS61;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 2.

```

```
DE Putative nitrite reductase (Fragment).
GN Name=nirk;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Throckack I.N., Pell M., Johansson M., Hallin S.;
RT "Denitrifying communities in soil are affected strongly by the heavy
   metal silver.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ304334; ABC02746.1; -; Genomic DNA.
DR EMBL; DQ304292; ABC02604.1; -; Genomic DNA.
FT NON_TER 1 158
FT NON_TER 1 158
SQ SEQUENCE 158 AA; 17375 MW; 192B3B2F5ED24C2A CRC64;

Query Match          55.6%; Score 5; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RETWF 5
Db 125 RETWF 129

RESULT 39
Q2PSC6_9BACT
ID Q2PSC6_9BACT PRELIMINARY; PRT; 158 AA.
AC Q2PSC6;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Putative nitrite reductase (Fragment).
GN Name=nirk;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Throckack I.N., Pell M., Johansson M., Hallin S.;
RT "Denitrifying communities in soil are affected strongly by the heavy
   metal silver.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ304369; ABC02681.1; -; Genomic DNA.
FT NON_TER 1 158
FT NON_TER 1 158
SQ SEQUENCE 158 AA; 17283 MW; 261AF19CECECB513 CRC64;

Query Match          55.6%; Score 5; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RETWF 5
Db 125 RETWF 129

RESULT 40
Q2PSG0_9BACT
ID Q2PSG0_9BACT PRELIMINARY; PRT; 158 AA.
AC Q2PSG0;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 2.
```

```
DE Putative nitrite reductase (Fragment).
GN Name=nirk;
OS uncultured bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=77133;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Throckack I.N., Pell M., Johansson M., Hallin S.;
RT "Denitrifying communities in soil are affected strongly by the heavy
   metal silver.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ304335; ABC02647.1; -; Genomic DNA.
FT NON_TER 1 158
FT NON_TER 1 158
SQ SEQUENCE 158 AA; 17349 MW; 44474B2F42BE3553 CRC64;

Query Match          55.6%; Score 5; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RETWF 5
Db 125 RETWF 129
```

Search completed: August 31, 2006, 10:39:28  
Job time : 149.25 secs

**THIS PAGE BLANK (USPTO)**



Result No.	Query	Score	Match	Length	DB	ID	Description
1	6	66.7	234	8	ADX89331	Adx89331	Plant full
2	6	66.7	316	8	ADX95024	Adx95024	Plant full
3	6	66.7	316	8	ADX94779	Adx94779	Plant full
4	6	66.7	392	6	AU061286	Au061286	Propionib
5	6	66.7	392	6	ABM57805	Abm57805	Propionib
6	6	66.7	484	6	ADT58120	Adt58120	Plant pol
7	6	66.7	494	7	ABM88426	Abm88426	Rice abio
8	5	55.6	7	2	AAV16490	Aay16490	Linear th
9	5	55.6	7	9	AEA36659	Aea36659	Thrombin
10	5	55.6	14	2	AAK47230	Aar47230	Thrombin
11	5	55.6	14	4	AAK23295	Aab23295	Hamster t
12	5	55.6	18	4	AAAG77283	Aag77283	Human col
13	5	55.6	30	4	ABB50666	Abb50666	Human sec
14	5	55.6	30	6	ABO44923	Ab044923	Novel hum
15	5	55.6	30	7	ABO26403	Ab026403	Protein a
16	5	55.6	35	3	AAV65136	Aay65136	Human 5'
17	5	55.6	35	8	ADU72700	Adu72700	Signal pe
18	5	55.6	35	9	ADZ73691	Adz73691	Human inc
19	5	55.6	36	2	AAW77907	Aar77907	Antigenic
20	5	55.6	36	2	AAW46100	Aaw46100	Predicted
21	5	55.6	36	2	AAV51705	Aay51705	H. influe
22	5	55.6	36	2	AAW53060	Aaw53060	Tbpi anti
23	5	55.6	36	3	AAV80402	Aay80402	H. influe

97	5	55.6	244	9	AEA27137	Stress to	KW	growth rate; cell cycle pathway; disease resistance;
98	5	55.6	254	2	AAR92511	VR-2332 O	KW	galactomannan production; lignin production; plant growth regulator;
99	5	55.6	254	2	AAR94720	PRRSV VR	KW	yield; plant growth; plant development; seed oil; protein yield;
100	5	55.6	254	2	AAW25975	ORF 3 pro	XX	protein content.
101	5	55.6	254	2	AAW25967	ORF 3 pro	XX	Unidentified.
102	5	55.6	254	2	AAW25955	ORF 3 pro	OS	
103	5	55.6	254	2	AAW25951	ORF 3 pro	XX	
104	5	55.6	254	2	AAW25959	ORF 3 pro	PN	
105	5	55.6	254	2	AAW25963	ORF 3 pro	XX	US2004034888-A1.
106	5	55.6	254	2	AAW55992	Porcine r	PD	19-FEB-2004.
107	5	55.6	254	2	ADG14054	Porcine r	XX	
108	5	55.6	254	3	AAY58671	Porcine r	PF	28-APR-2003; 2003US-00425114.
109	5	55.6	257	8	ADX66915	Plant ful	XX	06-MAY-1999; 99US-00304517.
110	5	55.6	263	9	AEC04333	Human bre	PR	05-NOV-2001; 2001US-00985678.
111	5	55.6	265	2	AAR29942	Deduced f	PR	
112	5	55.6	265	2	AAR74639	PRRS viru	XX	(LIU//) LIU J.
113	5	55.6	265	2	AAR88703	Porcine r	PA	(ZHOU//) ZHOU Y.
114	5	55.6	265	2	AAR94723	PRRSV Lel	PA	(KOVA//) KOVALIC D K.
115	5	55.6	265	4	AU071134	Porcine r	PA	(SCRE//) SCREEN S E.
116	5	55.6	265	5	AAU76136	Porcine r	PA	(TAB//) TABASKA J E.
117	5	55.6	265	5	ABG96499	PRRS viru	PA	(CAOY//) CAO Y.
118	5	55.6	265	5	ABG96505	PRRS viru	XX	
119	5	55.6	277	4	AAE03264	Human gen	XX	
120	5	55.6	278	8	ADQ25888	Human GPC	PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
121	5	55.6	307	7	ABO73671	Pseudomon	XX	WPI; 2004-180133/17.
122	5	55.6	308	8	ADN23722	Bacterial	DR	
123	5	55.6	310	3	AAG50071	Arabidops	XX	New recombinant DNA construct, useful for improving plant tolerance to
124	5	55.6	310	3	AGS21034	Arabidops	PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
125	5	55.6	315	5	ABP40601	Staphyloc	PT	pests, for conferring increased resistance to plant disease, or for
126	5	55.6	315	8	ADSO7806	Staphyloc	PT	improving yield.
127	5	55.6	324	3	AAB19419	A. preyit	XX	
128	5	55.6	324	3	AAE21033	Arabidops	XX	Claim 1; SEQ ID NO 51995; 15pp; English.
129	5	55.6	324	3	AAG50070	Arabidops	PS	
130	5	55.6	324	5	ABB81711	Synechocy	XX	The invention describes a recombinant DNA construct comprising a
131	5	55.6	324	5	AAU72781	Synechocy	CC	polynucleotide consisting of a sequence encoding an amino acid sequence
132	5	55.6	324	8	ADJ26726	Synechocy	CC	available in electronic form from the US patent office at
133	5	55.6	324	8	ADN20151	Bacterial	CC	ftp.segdata.upto.gov/sequence.html?docID:2004034888. The polynucleotide
134	5	55.6	328	6	ABU30747	Protein e	CC	of the invention are also useful in physical arrays of molecules and as
135	5	55.6	329	10	AEF28781	Lead Cere	CC	plant breeding markers. The recombinant DNA construct is useful for
136	5	55.6	331	8	ADJ71991	Human PWM	CC	improving plant tolerance to cold, heat, drought, herbicides, extreme
137	5	55.6	332	6	ABM70546	Phototrab	CC	osmotic conditions, pathogens or pests, for manipulating growth rate in
138	5	55.6	333	3	AAG41837	Arabidops	CC	plant cells by modification of the cell cycle pathway, for conferring
139	5	55.6	335	7	ADF06155	Bacterial	CC	increased resistance to plant disease, for producing galactomannan,
140	5	55.6	338	8	AAW98827	H. pylori	CC	lignin or plant growth regulators, for increasing the rate of homologous
141	5	55.6	343	8	ADT60944	Plant pol	CC	recombination in plants, for improving yield by modification of
142	5	55.6	347	8	ADN46788	Thermococ	CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
143	5	55.6	348	3	AAG41836	Arabidops	CC	or by providing improved plant growth and development under at least one
144	5	55.6	349	4	AAE61631	Protein k	CC	stress condition or for modifying seed oil or protein yield and/or
145	5	55.6	349	10	AEF41817	Soybean S	CC	content. This is the amino acid sequence of a plant full length insert
146	5	55.6	350	8	ADX78611	Plant ful	CC	polypeptide that can be used in the recombinant DNA construct of the
147	5	55.6	351	4	AAU32779	Novel hum	CC	invention.
148	5	55.6	354	8	ADY13581	Plant ful	XX	
149	5	55.6	355	8	ADX87488	Plant ful	XX	
150	5	55.6	359	3	AAG53574	Arabidops	SQ	Sequence 234 AA;

## ALIGNMENTS

Query Match 66.7%; Score 6; DB 8; Length 234;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLRN 8  
DB 127 SWFLRN 132  
|||||

RESULT 2  
ADX95024  
ID ADX95024 standard; protein; 314 AA.  
XX  
AC ADX95024;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Plant full length insert polypeptide seqid 57688.

```

XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
XX protein content.
XX Unidentified.
XX OS
XX US2004034888-A1.
XX PN
XX 19-FEB-2004.
XX PD
XX 28-APR-2003; 2003US-00425114.
XX PF
XX 06-MAY-1999; 99US-00304517.
XX PR
XX 05-NOV-2001; 2001US-00985678.
XX PR
XX (LIUJ/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALIC D K.
XX PA (SCRE/) SCREEN S E.
XX PA (TABA/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX DR
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 57688; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 314 AA;
SQ
Query Match 66.7%; Score 6; DB 8; Length 314;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SWFLRN 8
Db 152 SWFLRN 157
|||||
RESULT 3
ADX94779
ID ADX94779 standard; protein; 316 AA.
XX

```

```

AC ADX94779;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polypeptide seqid 57443.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
XX protein content.
XX Unidentified.
XX OS
XX US2004034888-A1.
XX PN
XX 19-FEB-2004.
XX PD
XX 28-APR-2003; 2003US-00425114.
XX PF
XX 06-MAY-1999; 99US-00304517.
XX PR
XX 05-NOV-2001; 2001US-00985678.
XX PR
XX (LIUJ/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALIC D K.
XX PA (SCRE/) SCREEN S E.
XX PA (TABA/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX DR
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 57443; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 316 AA;
SQ
Query Match 66.7%; Score 6; DB 8; Length 316;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SWFLRN 8
Db 154 SWFLRN 159
|||||

```

```
RESULT 4
AAU61286
ID AAU61286 standard; protein; 392 AA.
XX
AC AAU61286;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #22182.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
XX
PR 02-JUN-2000; 2000US-0208841P.
XX
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59616.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 22481; 1069pp; English.
XX
SS Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 392 AA;
Query Match 66.7%; Score 6; DB 4; Length 392;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VESWFL 6
DB 10 VESWFL 15
```

```
RESULT 5
ABM57805
ID ABM57805 standard; protein; 392 AA.
XX
AC ABM57805;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #22481.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieue-Douglas J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64545.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 22481; 1481pp; English.
XX
SS The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 392 AA;
Query Match 66.7%; Score 6; DB 6; Length 392;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 VESWFL 6  
 Db 10 VESWFL 15

RESULT 6  
 ADT58120 standard; protein; 484 AA.

XX ID ADT58120 standard; protein; 484 AA.  
 XX AC ADT58120;  
 XX XX  
 XX 13-JAN-2005 (first entry)  
 XX XX  
 XX Plant polypeptide, SEQ ID 8197.  
 XX XX  
 XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;  
 KW disease resistance; galactomannan production; plant growth regulator;  
 KW heat tolerance; herbicide tolerance; lignin production;  
 KW extreme osmotic condition tolerance; pathogens resistance;  
 KW pest resistance; yield improvement; seed oil yield; seed protein yield.  
 XX XX  
 XX Viridiplantae.  
 XX XX  
 XX US2004216190-A1.  
 XX XX  
 XX 28-OCT-2004.  
 XX XX  
 XX 18-DEC-2003; 2003US-00739930.  
 XX XX  
 XX 28-APR-2003; 2003US-00424599.  
 XX 28-APR-2003; 2003US-00425115.  
 XX XX  
 XX (KOVA/) KOVALIC D K.  
 XX XX  
 XX Kovalic DK;  
 XX XX  
 XX WPI; 2004-757369/74.  
 XX XX  
 XX New recombinant DNA constructs useful in the field of biochemistry and  
 PT genetics, and in particular for producing transgenic plants with improved  
 PT biological characteristics.  
 XX XX  
 XX Claim 2; SEQ ID NO 8197; 14pp; English.

The invention relates a recombinant DNA construct comprising a  
 CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:  
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences  
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,  
 CC Arabidopsis, wheat and rape but the specification does not indicate which  
 CC sequences is derived from which organism. Also included is a method of  
 CC producing a plant having an improved property, comprising transforming a  
 CC plant with a recombinant DNA construct comprising a promoter region  
 CC functional in a plant cell operably joined to a polynucleotide encoding a  
 CC polypeptide associated with the property, and growing the transformed  
 CC plant. The property is selected from improving plant cold tolerance, for  
 CC manipulating growth rate in plant cells by modification of the cell cycle  
 CC pathway, for improving plant drought tolerance, for providing increased  
 CC resistance to plant disease, for galactomannan production, for production  
 CC of plant growth regulators, for improving plant heat tolerance, for  
 CC improving plant tolerance to herbicides, for increasing the rate of  
 CC homologous recombination in plants, for lignin production, for improving  
 CC plant tolerance to extreme osmotic conditions, for improving plant  
 CC tolerance to pathogens or pests, for yield improvement by modification of  
 CC photosynthesis, for modifying seed oil yield and/or content, for  
 CC modifying seed protein yield and/or content, for yield improvement by  
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC and for yield improvement by providing improved plant growth and  
 CC development under at least one stress condition. The polynucleotide may  
 CC also encode a plant transcription factor. The methods and compositions of  
 CC the present invention are useful in the field of biochemistry and  
 CC genetics, in particular for producing transgenic plants with improved  
 CC biological characteristics such as increased yield, improved nitrogen  
 CC flow, increasing plant tolerance to cold or heat, improving plant

CC tolerance to extreme osmotic and drought conditions, and improving plant  
 CC tolerance to plant pests or pathogens. They can also be used in physical  
 CC arrays of molecules, plant breeding markers, computer-based storage and  
 CC analysis systems. The present sequence is one of the 5544 plant protein  
 CC sequences of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20040216190.

XX SQ Sequence 484 AA;

Query Match 66.7%; Score 6; DB 8; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 83;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SWFLRN 8  
 Db 287 SWFLRN 292

RESULT 7  
 ABM88426  
 ID ABM88426 standard; protein; 494 AA.  
 XX AC ABM88426;  
 XX XX  
 XX 02-JUN-2005 (first entry)  
 XX XX  
 XX Rice abiotic stress responsive polypeptide SEQ ID NO:6672.  
 XX DE  
 XX abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.  
 XX KW  
 XX Oryza sativa.  
 XX OS  
 XX WO2003008540-A2.  
 XX PN  
 XX 30-JAN-2003.  
 XX PD  
 XX 21-JUN-2002; 2002WO-US019668.  
 XX PF  
 XX 22-JUN-2001; 2001US-0300112P.  
 XX PR  
 XX 24-AUG-2001; 2001US-0314662P.  
 XX PR  
 XX 26-SEP-2001; 2001US-0325277P.  
 XX PR  
 XX 21-NOV-2001; 2001US-0332132P.  
 XX XX  
 XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX PA  
 XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;  
 XX PI Moughamer T, Provart N, Ricke D, Zhu T;  
 XX DR WPI; 2003-248011/24.  
 XX XX  
 XX New stress-responsive nucleic acid, useful for altering the  
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold  
 PT stress, salt stress or osmotic stress.  
 XX XX  
 XX Claim 1; SEQ ID NO 6672; 89pp; English.

The invention relates to novel abiotic stress responsive polynucleotides  
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host  
 CC cells, and plants containing such polynucleotides. Also disclosed are  
 CC methods for using the polynucleotides and polypeptides to alter the  
 CC responsiveness of a plant to abiotic stress. The invention is useful in  
 CC agriculture. The nucleic acid is useful for determining whether a test  
 CC plant has been exposed to an abiotic stress condition. It is also useful  
 CC for selecting an agent that alters abiotic stress regulated  
 CC polynucleotide expression in a plant cell, and to identify a homolog or  
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid  
 CC molecule and the polypeptide encoded by it are useful in altering the  
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt  
 CC stress, osmotic stress or any of their combinations. The present sequence  
 CC is used in the exemplification of the invention  
 XX XX





AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patients own production of P. Additionally, N may be used to produce the colon cancer-associated Ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent sequences used in the exemplification of the present invention. N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922

XX Sequence 18 AA;

Query Match 55.6%; Score 5; DB 4; Length 18;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ESWFL 6

Db 9 ESWFL 13

RESULT 13

ABBS0666

ID ABBS0666 standard; protein; 30 AA.

AC ABBS0666;

DT 07-FEB-2002 (first entry)

DE Human secreted protein encoded by gene 44 SEQ ID NO:614.

Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV; dermatological; immunosuppressive; antiinflammatory; immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic; ophthalmological; neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnerary; antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder; multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer; human immunodeficiency virus; hyperproliferative disorder; wound healing; Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis; Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder; corneal graft neovascularisation; diabetic retinopathy; regeneration; neurological disorder; Huntington's chorea; Alzheimer's disease; Parkinson's disease; infectious disease.

XX Homo sapiens.

OS WO200162891-A2.

PN 30-AUG-2001.

XX 21-FEB-2001; 2001WO-US005614.

XX 24-FEB-2000; 2000US-0184836P.

PR 29-MAR-2000; 2000US-0193170P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Ebner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;

PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;

PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrie AM, Fan P;

PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G, Zeng Z;

PI Greene JM;

XX WPI; 2001-625724/72.

XX Nucleic acids encoding 207 human secreted polypeptides, useful for preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease and diabetic retinopathy.

PS Disclosure; Page 101; 1533pp; English.

XX ABBS0301 to ABBS1287 and ABA83194 to ABA83441 represent human secreted proteins (I) and polynucleotide (II) sequences. (I) and (II) have various activities based on the tissues and cells the genes are expressed in. Example of these activities include: immunomodulatory; antisclerotic; dermatological; immunosuppressive; antiinflammatory; immunostimulant; anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological; neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular; antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used in gene therapy and vaccine production. (I) and (II) can be used in the prevention, diagnosis and treatment of immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus and human immunodeficiency virus (HIV) infections), hyperproliferative disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft neovascularisation and diabetic retinopathy), neurological disorders (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease), infectious diseases and/or for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to ABA83193 and ABBS0300 represent sequences used in the exemplification of the present invention

XX Sequence 30 AA;

Query Match 55.6%; Score 5; DB 4; Length 30;

Best Local Similarity 100.0%; Pred. No. 99;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9

Db 1 FLRNP 5

RESULT 14

ABO44923

ID ABO44923 standard; protein; 30 AA.

AC ABO44923;

DT 02-OCT-2003 (first entry)

XX Novel human secreted protein #44 fragment #2.

Human; gene therapy; autoimmune disorder; multiple sclerosis; cancer; systemic lupus erythematosus; haematopoietic cell disorder; allergy; agammaglobulinaemia; ataxia telangiectasia; blood coagulation disorder; afibrinogenemia; thrombocytopenia; graft-versus-host disease; arthritis; inflammatory condition; ischaemia-reperfusion injury; infectious disease; hyperproliferative disorder; purpura; viral infection; regeneration; bacterial infection; ulcer; Alzheimer's disease.

XX Homo sapiens.

OS US2003065160-A1.

PN 03-APR-2003.

XX 07-DEC-2001; 2001US-00004860.

XX 06-JUN-1997; 97US-0048875P.

PR 06-JUN-1997; 97US-0048876P.

PR 06-JUN-1997; 97US-0048877P.

PR 06-JUN-1997; 97US-0048878P.

PR 06-JUN-1997; 97US-0048880P.

PR 06-JUN-1997; 97US-0048881P.

PR 06-JUN-1997; 97US-0048882P.

PR 06-JUN-1997; 97US-0048883P.





PR 06-JUN-1997; 97US-0048884P.  
 PR 06-JUN-1997; 97US-0048885P.  
 PR 06-JUN-1997; 97US-0048892P.  
 PR 06-JUN-1997; 97US-0048893P.  
 PR 06-JUN-1997; 97US-0048894P.  
 PR 06-JUN-1997; 97US-0048895P.  
 PR 06-JUN-1997; 97US-0048896P.  
 PR 06-JUN-1997; 97US-0048897P.  
 PR 06-JUN-1997; 97US-0048898P.  
 PR 06-JUN-1997; 97US-0048899P.  
 PR 06-JUN-1997; 97US-0048900P.  
 PR 06-JUN-1997; 97US-0048901P.  
 PR 06-JUN-1997; 97US-0048915P.  
 PR 06-JUN-1997; 97US-0048916P.  
 PR 06-JUN-1997; 97US-0048917P.  
 PR 06-JUN-1997; 97US-0048949P.  
 PR 06-JUN-1997; 97US-0048962P.  
 PR 06-JUN-1997; 97US-0048963P.  
 PR 06-JUN-1997; 97US-0048964P.  
 PR 06-JUN-1997; 97US-0048970P.  
 PR 06-JUN-1997; 97US-0048971P.  
 PR 06-JUN-1997; 97US-0048972P.  
 PR 06-JUN-1997; 97US-0048974P.  
 PR 06-JUN-1997; 97US-0049019P.  
 PR 06-JUN-1997; 97US-0049020P.  
 PR 06-JUN-1997; 97US-0049373P.  
 PR 06-JUN-1997; 97US-0049374P.  
 PR 06-JUN-1997; 97US-0049375P.  
 PR 05-SEP-1997; 97US-0057584P.  
 PR 05-SEP-1997; 97US-0057627P.  
 PR 05-SEP-1997; 97US-0057628P.  
 PR 05-SEP-1997; 97US-0057629P.  
 PR 05-SEP-1997; 97US-0057634P.  
 PR 05-SEP-1997; 97US-0057635P.  
 PR 05-SEP-1997; 97US-0057642P.  
 PR 05-SEP-1997; 97US-0057643P.  
 PR 05-SEP-1997; 97US-0057644P.  
 PR 05-SEP-1997; 97US-0057645P.  
 PR 05-SEP-1997; 97US-0057646P.  
 PR 05-SEP-1997; 97US-0057647P.  
 PR 05-SEP-1997; 97US-0057648P.  
 PR 05-SEP-1997; 97US-0057649P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057651P.  
 PR 05-SEP-1997; 97US-0057654P.  
 PR 05-SEP-1997; 97US-0057661P.  
 PR 05-SEP-1997; 97US-0057662P.  
 PR 05-SEP-1997; 97US-0057666P.  
 PR 05-SEP-1997; 97US-0057667P.  
 PR 05-SEP-1997; 97US-0057668P.  
 PR 05-SEP-1997; 97US-0057760P.  
 PR 05-SEP-1997; 97US-0057761P.  
 PR 05-SEP-1997; 97US-0057762P.  
 PR 05-SEP-1997; 97US-0057763P.  
 PR 05-SEP-1997; 97US-0057764P.  
 PR 05-SEP-1997; 97US-0057765P.  
 PR 05-SEP-1997; 97US-0057769P.  
 PR 05-SEP-1997; 97US-0057770P.  
 PR 05-SEP-1997; 97US-0057771P.  
 PR 05-SEP-1997; 97US-0057774P.  
 PR 05-SEP-1997; 97US-0057775P.  
 PR 05-SEP-1997; 97US-0057776P.  
 PR 05-SEP-1997; 97US-0057777P.  
 PR 05-SEP-1997; 97US-0057778P.  
 PR 18-DEC-1997; 97US-0070923P.  
 PR 04-JUN-1998; 98WO-US011422.  
 PR 15-JUL-1998; 98US-0092921P.  
 PR 30-JUL-1998; 98US-0094657P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;  
 PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;  
 PI  
 97US-0048884P.  
 97US-0048885P.  
 97US-0048892P.  
 97US-0048893P.  
 97US-0048894P.  
 97US-0048895P.  
 97US-0048896P.  
 97US-0048897P.  
 97US-0048898P.  
 97US-0048899P.  
 97US-0048900P.  
 97US-0048901P.  
 97US-0048915P.  
 97US-0048916P.  
 97US-0048917P.  
 97US-0048949P.  
 97US-0048962P.  
 97US-0048963P.  
 97US-0048964P.  
 97US-0048970P.  
 97US-0048971P.  
 97US-0048972P.  
 97US-0048974P.  
 97US-0049019P.  
 97US-0049020P.  
 97US-0049373P.  
 97US-0049374P.  
 97US-0049375P.  
 97US-0057584P.  
 97US-0057627P.  
 97US-0057628P.  
 97US-0057629P.  
 97US-0057634P.  
 97US-0057635P.  
 97US-0057642P.  
 97US-0057643P.  
 97US-0057644P.  
 97US-0057645P.  
 97US-0057646P.  
 97US-0057647P.  
 97US-0057648P.  
 97US-0057649P.  
 97US-0057650P.  
 97US-0057651P.  
 97US-0057654P.  
 97US-0057661P.  
 97US-0057662P.  
 97US-0057666P.  
 97US-0057667P.  
 97US-0057668P.  
 97US-0057760P.  
 97US-0057761P.  
 97US-0057762P.  
 97US-0057763P.  
 97US-0057764P.  
 97US-0057765P.  
 97US-0057769P.  
 97US-0057770P.  
 97US-0057771P.  
 97US-0057774P.  
 97US-0057775P.  
 97US-0057776P.  
 97US-0057777P.  
 97US-0057778P.  
 97US-0070923P.  
 98WO-US011422.  
 98US-0092921P.  
 98US-0094657P.  
 (GEST ) GENSET.  
 Dumas Milne Edwards J, Duclert A, Giordano J;  
 WPI: 2000-038446/03.  
 N-PSDB; AA242750.  
 Novel secreted protein 5' expressed sequence tag sequences used in

PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;  
 PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;  
 PI Carter KC;  
 XX WPI: 2003-511926/48.  
 DR  
 XX New precerebellin-like protein, useful for diagnosing or treating  
 PT neurodegenerative and behavioral disorders, immune disorders, liver  
 PT disorders, and cancer.  
 XX  
 PS Disclosure; Col 48; 156pp; English.  
 XX  
 CC The invention relates to an isolated protein comprising amino acid  
 CC residues 33-205 or 1-205 of a novel human secreted protein appearing as  
 CC ABO26252. The protein is encoded by one of 238 disclosed cDNA sequences  
 CC encoding 238 secreted proteins. ABO26252 is a precerebellin-like protein.  
 CC Also included are a composition comprising the protein and a carrier and  
 CC an isolated protein produced by expressing the protein cited above by a  
 CC cell, and recovering the protein. The proteins are useful for diagnosing  
 CC or treating neurodegenerative and behavioural disorders (e.g. Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease, schizophrenia, mania,  
 CC dementia, paranoia, psychoses or autism), immune disorders (e.g.  
 CC infection, inflammation, allergy), liver disorders (e.g. hepatoblastoma,  
 CC jaundice, hepatitis), immunological disorders (e.g. AIDS, leukaemia,  
 CC rheumatoid arthritis, sepsis, acne, psoriasis) and cancer. The present  
 CC sequence is a protein associated with one of the 238 disclosed novel  
 CC secreted proteins  
 XX  
 SQ Sequence 30 AA;  
 Query Match 55.6%; Score 5; DB 7; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 99;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRNP 9  
 Db |||||  
 1 FLRNP 5  
 RESULT 16  
 AAY65136  
 ID AAY65136 standard; protein; 35 AA.  
 XX  
 AC AAY65136;  
 XX  
 DT 01-FEB-2000 (first entry)  
 XX  
 DE Human 5' EST related polypeptide SEQ ID NO:1297.  
 XX  
 KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;  
 KW gene therapy; chromosome mapping; upstream regulatory sequence; forensic;  
 KW location; development; protein synthesis; stability; regulation;  
 KW identification.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9593051-A2.  
 XX  
 PD 21-OCT-1999.  
 XX  
 PF 09-APR-1999; 99WO-IB000712.  
 XX  
 PR 09-APR-1998; 98US-00057719.  
 PR 28-APR-1998; 98US-00069047.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Dumas Milne Edwards J, Duclert A, Giordano J;  
 DR WPI: 2000-038446/03.  
 DR N-PSDB; AA242750.  
 XX  
 PT Novel secreted protein 5' expressed sequence tag sequences used in

PT diagnostic, forensic, gene therapy, and chromosome mapping procedures.  
PS Claim 3; Page 744; 837pp; English.  
XX  
CC AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)  
CC sequences, corresponding to human secreted proteins. AA242265 to AA243075  
CC represent the EST-related proteins corresponding to AA242265 to AA243075.  
CC The 5' ESTs can be used for producing secreted human gene products. They  
CC can be used to identify and isolate 5' untranslated regions (UTRs) and  
CC upstream regulatory regions which control the location, development  
CC stage, rate, and quantity of protein synthesis, as well as stability of  
CC mRNA. The ESTs are also useful as probes for chromosome mapping, and to  
CC obtain full length cDNA clones. The ESTs can also be used in forensic  
CC procedures to identify individuals, or in diagnostic procedures to  
CC identify individuals having genetic diseases resulting from abnormal gene  
CC expression. The products may also be used in gene therapy protocols. The  
CC nucleic acids encoding signal peptides can be used for directing  
CC extracellular secretion of a polypeptide or the insertion of a  
CC polypeptide into a membrane, or importing a polypeptide into a cell. The  
CC proteins encoded by the EST sequences may be useful in treating a variety  
CC of human conditions. Secreted proteins have therapeutic value, and the  
CC identification of new secreted proteins is valuable. AA242249 to AA242264  
CC and AA242264 to AA242265 represent sequences used in the exemplification  
CC of the present invention  
XX  
SQ Sequence 35 AA;  
Query Match 55.6%; Score 5; DB 3; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 SWFLR 7  
DB 9 SWFLR 13  
RESULT 17  
ADU72700  
ID ADU72700 standard; peptide; 35 AA.  
XX  
AC ADU72700;  
XX  
DT 10-FEB-2005 (first entry)  
XX  
XX Signal peptide-containing polypeptide fragment, SEQ ID NO:1297.  
DE Protein secretion; recombinant protein; diagnosis; mapping; forensic;  
XX gene therapy; signal peptide.  
KW  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..22  
FT /label= Signal\_peptide  
FT Protein 23..35  
FT /label= Mature\_polypeptide\_fragment  
XX  
XX US6822072-B1.  
XX  
XX 23-NOV-2004.  
XX  
XX 21-DEC-1999; 99US-00471276.  
XX  
XX 09-APR-1998; 98US-00057719.  
XX 28-APR-1998; 98US-00069047.  
XX 09-APR-1999; 99WO-IB000712.  
XX (GENT ) GENSET SA.  
XX Edwards JDM, Duclert A, Giordano J;  
XX WFI; 2004-812112/80.  
DR N-PSDB; ADU71912.

XX New expressed sequence tags and encoded human proteins useful for  
PT diagnosing, preventing or treating diseases such as autoimmune disorders,  
PT inflammation, wounds or infections, or in forensic or chromosome mapping  
PT procedures.  
XX Example 15; SEQ ID NO 1297; 72pp; English.  
XX The invention relates to an isolated or purified signal peptide  
CC consisting of residues 1-16 of ADU72234 (signal peptide given separately  
CC as ADU73026) which directs the extracellular secretion of a polypeptide  
CC to which it is operably linked. The invention also relates to a method of  
CC producing the signal peptide. The invention further discloses: isolated,  
CC purified or enriched 5' expressed sequence tags (ESTs), many of which  
CC encode all or a part of a secretory signal peptide; polypeptides encoded  
CC by these ESTs (EST-related polypeptides); antibodies which recognize the  
CC EST-related polypeptides; vectors and host cells comprising EST-related  
CC nucleic acids of the invention; an array of ESTs; methods involving the  
CC use of signal peptides of the invention to target polypeptides; and  
CC methods involving the use of ESTs of the invention, for example, in  
CC identifying a promoter in genomic DNA. The EST-encoded signal peptides of  
CC the invention are useful for directing the secretion or import of a  
CC recombinant polypeptide via the generation of protein fusions comprising  
CC such signal peptides. The ESTs, EST-related polypeptides and methods of  
CC the invention can be used for forensic procedures, chromosome mapping,  
CC diagnostics, and therapeutic procedures, including gene therapy.  
CC Sequences ADU72215-ADU72919 represent incomplete polypeptides comprising  
CC a signal peptide which are encoded by the 5' ESTs shown in ADU71427-  
CC ADU72131. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from the US patent office at  
CC seqdata.uspto.gov/sequence.html?DocID=US6822072.  
XX  
SQ Sequence 35 AA;  
Query Match 55.6%; Score 5; DB 8; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 SWFLR 7  
DB 9 SWFLR 13  
RESULT 18  
ADZ73691  
ID ADZ73691 standard; peptide; 35 AA.  
XX  
AC ADZ73691;  
XX  
DT 28-JUL-2005 (first entry)  
XX  
XX Human incomplete polypeptide including a signal peptide SEQ ID NO:1297.  
DE expressed sequence tag; EST; expression; protein secretion; diagnostic;  
XX forensic; gene therapy; haplotype mapping.  
XX Homo sapiens.  
XX  
XX US2005106595-A1.  
XX  
XX 19-MAY-2005.  
XX  
XX 25-AUG-2004; 2004US-00926683.  
XX  
XX 09-APR-1998; 98US-00057719.  
XX 28-APR-1998; 98US-00069047.  
XX 09-APR-1999; 99WO-IB000712.  
XX 21-DEC-1999; 99US-00471276.  
XX  
XX (GENT ) GENSET SA.  
XX  
XX Dumas MEJ, Duclert A, Giordano J;

XX WPI: 2005-384300/39.  
 DR N-PSDB; ADZ72903.  
 XX New purified nucleic acid expressing secreted proteins useful in  
 PT forensic, gene therapy, and chromosome mapping procedures, and diagnosing  
 PT or treating cancer, atherosclerosis and autoimmune diseases, diabetes,  
 PT asthma and infections.  
 XX Claim 1; SEQ ID NO 1297; 79pp; English.  
 XX The invention relates to a novel purified nucleic acid (I) comprising any  
 CC of (ADZ72418-ADZ73205) or (ADZ73994-ADZ74016) and their complements; at  
 CC least 15 consecutive nucleotides of (I) and their complements; or any of  
 CC 788 nucleotide sequences encoding fully defined sequences of 16-255 amino  
 CC acids (ADZ73206-ADZ73993). The invention discloses 5' EST's derived from  
 CC mRNAs encoding secreted proteins. The 5' EST's may be used to obtain  
 CC cDNAs and genomic DNAs corresponding to the 5' ESTs. The methods and  
 CC compositions of the present invention are useful for expressing secreted  
 CC proteins or its portions (claimed) or to obtain antibodies capable of  
 CC specifically binding to the secreted proteins, and in diagnostic,  
 CC forensic, gene therapy, and chromosome mapping procedures, and for  
 CC designing expression vectors and secretion vectors. The present sequence  
 CC is used in the exemplification of the invention. Note: The sequence data  
 CC for this patent is not represented in the printed specification but is  
 CC based on sequence information supplied in electronic format from the  
 CC USPTO web site seqdata.uspto.gov/sequence.html; Document ID: 20050106595.  
 XX SQ Sequence 35 AA;

Query Match 55.6%; Score 5; DB 9; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SWFLR 7  
 Db |||||  
 9 SWFLR 13

## RESULT 19

AAR77907  
 ID AAR77907 standard; peptide; 36 AA.

XX AAR77907;

XX 09-OCT-1996 (first entry)

XX Antigenic Tbp1 peptide TBp1-11.

DE Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;  
 KW non-typable strain; Haemophilus influenzae; meningitis.

XX Synthetic.

XX WO9513370-A1.

XX 18-MAY-1995.

XX 07-NOV-1994; 94WO-CA000616.

XX 08-NOV-1993; 93US-00148968.

PR 29-DEC-1993; 93US-00175116.

XX (CONN-) CONNAUGHT LAB LTD.

PI Loosmore S, Harkness R, Schryvers A, Chong P, Gray-Owen S;  
 PI Yang Y, Murrin A, Klein M;

XX WPI; 1995-194089/25.

XX Nucleic acids encoding Haemophilus transferrin receptor - used to develop  
 PT prods for detection and in diagnosis, prevention and treatment of  
 PT Haemophilus infection.

XX Example 16; Page 70; 231pp; English.  
 PS AAR77897-932 are predicted antigenic peptides derived from conserved  
 CC regions of the Tbp1 protein from H. influenzae strains Egan, Minna, DL63  
 CC and non-typable strain PAK12085. The transferrin receptor (TfR) operon  
 CC consists of two genes (Tbp1 and Tbp2) arranged in tandem and which are  
 CC transcribed from a single promoter. H. influenzae TfR is iron- and/or  
 CC haemin-regulated and a putative fur-binding site has been identified  
 CC upstream of Tbp2. Antibodies blocking this binding site may prevent  
 CC bacterial growth. Fragments of the TfR (or its genes) are useful in  
 CC vaccines to provide protection against, e.g. bacterial meningitis. An  
 CC advantage of using the TfR is that it shares homology with TfR of other  
 CC H. influenzae strains including non-typable strains. The present sequence  
 CC shows residues 315-350 of Tbp1 from H. influenzae type b strain DL63 (see  
 CC AAR77894), not residues 293-328 as stated in the specification. Numbering  
 CC of the residues is said to be according to the H. influenzae strain Egan  
 CC (see AAR77886), but appears to be as for strain DL63

XX Sequence 36 AA;

Query Match 55.6%; Score 5; DB 2; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SWFLR 7  
 Db |||||

5 SWFLR 9

## RESULT 20

AAW46100

ID AAW46100 standard; protein; 36 AA.

XX AAW46100;

XX 05-MAY-1998 (first entry)

XX Predicted antigenic Tbp1 peptide TBp1-11.

DE Transferrin receptor; Haemophilus influenzae type b; iron;  
 KW human transferrin; iron source; antibody; bacterial growth; vaccine;  
 KW immunogenic truncated analogue; antigen; Tbp1; Tbp2.

XX Synthetic.

OS Haemophilus influenzae.

XX WO9640929-A2.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-CA000399.

XX 07-JUN-1995; 95US-00483577.

PR 17-MAY-1996; 96US-00649518.

XX (CONN-) CONNAUGHT LAB LTD.

PI Loosmore SM, Harkness RE, Schryvers AB, Chong P, Gray-Owen S;  
 PI Yang Y, Murrin AD, Klein MH;

XX WPI; 1997-052329/05.

PT Haemophilus truncated transferrin receptor protein analogue, Tbp2 - used  
 PT to induce protection against disease caused by transferrin producing  
 PT pathogens, or as antigen to detect Haemophilus TfR antibodies.

XX Example 16; Page 68; 228pp; English.

XX AAW46090-125 are predicted antigenic peptides derived from the Tbp1

CC protein of Haemophilus influenzae type b. Tbp1 is part of the transferrin  
 CC receptor, of which Tbp2 is also a subunit. The deduced amino acid  
 CC sequences of Tbp1 and Tbp2 were compared, and regions of conservation

CC identified. The above peptides are derived from these regions, the  
 CC present peptide being derived from residues 293-328. Iron is an essential  
 CC nutrient for the growth of these bacteria, and they can utilise human  
 CC transferrin as a source of iron. Antibodies which block the access of the  
 CC transferrin receptor to its iron source prevent bacterial growth. The  
 CC transferrin receptor, or fragments, therefore, are good vaccine  
 CC candidates. An immunogenic composition comprising (or encoding) the  
 CC immunogenic truncated analogue can be used to induce protection against a  
 CC disease caused by a bacterial pathogen that produces the transferrin  
 CC receptor. The immunogenic truncated analogue is also useful as an antigen  
 CC in immunoassays for the detection of Haemophilus transferrin receptor  
 CC antibodies, while the nucleic acid molecule can be used as a  
 CC hybridisation probe for the detection of other transferrin receptor genes  
 XX  
 SQ Sequence 36 AA;

Query Match 55.6%; Score 5; DB 2; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SWFLR 7  
 Db 5 SWFLR 9  
 |||||

RESULT 21  
 AAY51705  
 ID AAY51705 standard; protein; 36 AA.  
 XX  
 AC AAY51705;  
 XX  
 DT 13-JUN-2000 (first entry)  
 XX  
 DE H. influenzae antigenic Tbp1 peptide TBP1-11.  
 XX  
 KW Transferrin receptor; Tbp1; Tbp2; immunogenic; antibacterial; vaccine;  
 KW diagnosis.

XX Haemophilus influenzae.  
 XX US6015688-A.  
 XX  
 PD 18-JAN-2000.  
 XX  
 PF 07-JUN-1995; 95US-00483577.  
 XX  
 PR 08-NOV-1993; 93US-00148968.  
 PR 29-DEC-1993; 93US-00175116.  
 PR 08-NOV-1994; 94US-00337483.  
 XX  
 PA (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PI Loosmore S, Harkness R, Schryvers A, Gray-Owen S, Yang Y;  
 PI Murdin A, Klein M, Chong P;  
 XX  
 DR WPI; 1997-052329/05.  
 XX  
 PT Haemophilus truncated transferrin receptor protein analogue, Tbp2 - used  
 PT to induce protection against disease caused by transferrin producing  
 PT pathogens, or as antigen to detect Haemophilus TIR antibodies.  
 XX

PS Example 16; Col 37-38; 281pp; English.  
 XX  
 SS This invention describes a novel isolated and purified nucleic acid (I)  
 CC encoding an immunogenic, C-terminally truncated analog of one of the  
 CC transferrin receptor proteins Tbp1 or Tbp2 of Haemophilus influenzae  
 CC which has antibacterial activity. (I) are used for recombinant production  
 CC of truncated Tbp; as probes and primers for detecting, and diagnosing  
 CC infection by, Haemophilus, also for isolating similar sequences from  
 CC other bacteria; as immunogens for vaccinating against infections caused  
 CC by bacteria that produce transferrin receptors, e.g. Haemophilus,  
 CC Neisseria or Branhamella. The truncated proteins are useful as immunogens  
 CC (as above); for diagnosing infection (as antigens in immunoassays) and

CC for raising antibodies, used for diagnosis of infections or for passive  
 CC immunization. AAY51695-Y51767 represent H. influenzae transferrin  
 CC receptor proteins Tbp1 and Tbp2 antigenic peptide fragments  
 XX  
 SQ Sequence 36 AA;

Query Match 55.6%; Score 5; DB 2; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SWFLR 7  
 Db 5 SWFLR 9  
 |||||

RESULT 22  
 AAW53060  
 ID AAW53060 standard; peptide; 36 AA.  
 XX  
 AC AAW53060;  
 XX  
 DT 20-JUL-1998 (first entry)  
 XX  
 DE Tbp1 antigenic peptide TBP1-11.  
 XX  
 KW tbp1; tbp2; vaccine; H. influenzae; antibody; diagnosis;  
 KW passive immunisation; transferrin receptor operon.  
 XX  
 OS Haemophilus influenzae.  
 XX  
 PN US5708149-A.  
 XX  
 PD 13-JAN-1998.  
 XX  
 PF 07-JUN-1995; 95US-00487890.  
 XX  
 PR 08-NOV-1993; 93US-00148968.  
 PR 29-DEC-1993; 93US-00175116.  
 PR 08-NOV-1994; 94US-00337483.  
 XX  
 PA (CONN-) CONNAUGHT LAB LTD.  
 XX  
 PI Gray-Owen S, Klein M, Harkness R, Loosmore S, Yang Y, Chong P;  
 PI Murdin A, Schryvers A;  
 XX  
 DR WPI; 1998-100410/09.  
 XX  
 PT Purification of recombinant Haemophilus transferrin-binding protein - by  
 PT solubilising inclusion bodies separated from cell lysate.

PS Example 16/17; Column 35-36; 261pp; English.  
 XX  
 SS Peptides AAW53050-W53085 are derived from the Tbp1 protein. The Tbp1  
 CC protein is one of two proteins with genes found on the transferrin  
 CC operon. These peptides can be used along with the genes, DNA sequences  
 CC and recombinant proteins for diagnosis, immunisation and the generation  
 CC of diagnostic and immunological reagents. They can also be used to  
 CC protect from bacteria that produce transferrin receptor protein  
 XX

Query Match 55.6%; Score 5; DB 2; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SWFLR 7  
 Db 5 SWFLR 9  
 |||||

RESULT 23  
 AAY80402  
 ID AAY80402 standard; peptide; 36 AA.

```

XX AC AAY60402;
XX DT 06-JUN-2000 (first entry)
XX DE H. influenzae transferrin receptor Tbp1 epitope TBp1-11.
XX DE Antibacterial; antiinflammatory; auditory; respiratory; antibody;
KW antiserum; transferrin receptor; immunogen; epitope; otitis media;
KW bacterial meningitis; epiglottitis; pneumonia; tracheobronchitis.
XX OS Haemophilus influenzae.
XX PN US6008326-A.
XX PD 28-DEC-1999.
XX PF 07-JUN-1995; 95US-00474671.
XX PR 08-NOV-1993; 93US-00148968.
XX PR 29-DEC-1993; 93US-00175116.
XX PR 08-NOV-1995; 95US-00337483.
XX PA (CONN-) CONNAUGHT LAB LTD.
XX PI Loemore S, Harkness R, Chong P, Gray-Owen S, Yang Y, Klein M;
XX PI Murdin A, Schryvers A;
XX DR WPI; 2000-096387/08.
XX PT Antibodies specific for transferrin receptor proteins of Haemophilus
XX PT influenzae, useful for treating otitis media, epiglottitis, pneumonia and
XX PT tracheobronchitis.
XX PS Disclosure; Col 37-38; 252pp; English.
XX CC The invention relates to novel antibodies (or monospecific antisera)
XX CC specific for single transferrin receptor proteins (or immunogenic
XX CC fragment) from strains of Haemophilus influenzae. This sequence
XX CC corresponds to an epitope from the H. influenzae transferrin receptor
XX CC protein Tbp1. The antibodies may be used for preventing and treating
XX CC infections and disorders caused by H. influenzae, including bacterial
XX CC meningitis, otitis media, epiglottitis, pneumonia and tracheobronchitis.
XX CC The antibodies may also be used to detect the presence of H. influenzae
XX CC proteins in samples according to standard methodologies (e.g. enzyme
XX CC linked immunosorbent assay (ELISA)) and hence diagnose infections
XX SQ Sequence 36 AA;
Query Match 55.6%; Score 5; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SWFLR 7
Db 5 SWFLR 9
|||||
RESULT 24
AAU53548
ID AAU53548 standard; protein; 50 AA.
XX AC AAU53548;
XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #14444.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX OS
XX

```

```

OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US012865.
XX PR 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR WPI; 2001-616774/71.
XX DR N-PSDB; AAS59561.
XX PR Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX PS Example 1; SEQ ID NO 14743; 1069pp; English.
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 50 AA;
Query Match 55.6%; Score 5; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ESWFL 6
Db 44 ESWFL 48
|||||
RESULT 25
ABM50067
ID ABM50067 standard; protein; 50 AA.
XX AC ABM50067;
XX DT 20-OCT-2003 (first entry)
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #14743.
XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX KW immunostimulant; immune response; vaccine.
XX OS Propionibacterium acnes.
XX

```

```

PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persaing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieue-Douglas J;
XX
XX WPI; 2003-381789/36.
XX
XX N-PSDB; ACF64490.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 14743; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 50 AA;

Query Match 55.6%; Score 5; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ESWFL 6
DB 44 ESWFL 48

RESULT 26
AAM21922
ID AAM21922 standard; protein; 60 AA.
XX
XX AAM21922;
XX
XX 12-OCT-2001 (first entry)
XX
DE Peptide #8356 encoded by probe for measuring cervical gene expression.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KW

cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000670.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 28-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
XX Claim 27; SEQ ID NO 26748; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 60 AA;

Query Match 55.6%; Score 5; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7
DB 6 SWFLR 10

RESULT 27
ABB44299
ID ABB44299 standard; peptide; 60 AA.
XX
XX ABB44299;
XX
XX 04-FEB-2002 (first entry)
XX
XX Peptide #11805 encoded by human foetal liver single exon probe.
DE
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR

```

PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-483447/52.  
 DR  
 DR Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human fetal liver.  
 PT  
 PS Claim 27; SEQ ID NO 36934; 639pp + Sequence Listing; English.  
 XX  
 XX The invention relates to a single exon nucleic acid probe for measuring  
 CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human fetal liver. The  
 CC present sequence is a peptide encoded by a single exon nucleic acid probe  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 60 AA;  
 SQ  
 Query Match 55.6%; Score 5; DB 4; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SWFLR 7  
 DB |||||  
 6 SWFLR 10  
 RESULT 28  
 AAM38250  
 ID AAM38250 standard; protein; 60 AA.  
 XX  
 AC AAM38250;  
 DT 17-OCT-2001 (first entry)  
 XX  
 XX Peptide #12287 encoded by probe for measuring placental gene expression.  
 DE  
 XX Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157272-A2.  
 PN  
 XX  
 XX 09-AUG-2001.  
 PD  
 XX  
 PF 30-JAN-2001; 2001WO-US000663.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488897/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human placenta.  
 XX  
 PS Claim 27; SEQ ID NO 38519; 654pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SNP:  
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders  
 XX  
 XX Sequence 60 AA;  
 SQ  
 Query Match 55.6%; Score 5; DB 4; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SWFLR 7  
 DB |||||  
 6 SWFLR 10  
 RESULT 29  
 ABB27159  
 ID ABB27159 standard; protein; 60 AA.  
 XX  
 AC ABB27159;  
 XX  
 XX 23-JAN-2002 (first entry)  
 DT  
 XX Protein #9158 encoded by probe for measuring heart cell gene expression.  
 DE  
 XX Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157274-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX  
 PF 30-JAN-2001; 2001WO-US000666.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488899/53.  
 XX  
 DR Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts.  
 PT  
 PS Claim 15; SEQ ID NO 28929; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
 CC sequence data for this patent did not form part of the printed



RESULT 31

```

PN WO200157273-A2.
XX
PD
XX
XX
PF 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US0000664.
XX
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
DR WPI; 2001-488898/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
XX
PS Claim 27; SEQ ID NO 38316; 658pp; English.
XX
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (for complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 60 AA;
Query Match 55.6%; Score 5; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SWFLR 7
DB 6 SWFLR 10
RESULT 33
ABG47046
ID ABG47046 standard; peptide; 60 AA.
XX
XX
AC ABG47046;
XX
XX
DT 19-AUG-2002 (first entry)
XX
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 36711.
XX
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioloeyomytosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
XX
OS Homo sapiens.
XX

```

---

```

PN WO200186003-A2.
XX
XX
PD
XX
XX
PF 15-NOV-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US0000655.
XX
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
DR WPI; 2002-114183/15.
XX
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
XX
PS Claim 27; SEQ ID NO 36711; 634pp; English.
XX
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioloeyomytosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 60 AA;
Query Match 55.6%; Score 5; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SWFLR 7
DB 6 SWFLR 10

```

RESULT 34  
ABP07914  
XX ABP07914 standard; protein; 67 AA.  
XX  
XX  
XX ABP07914;  
XX  
XX  
XX 24-JUN-2002 (first entry)  
XX  
XX Human ORFX protein sequence SEQ ID NO:15810.  
XX  
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;  
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
XX hypertension; hypothyroidism; cholesterol ester storage disease;  
XX immune deficiency; immune disorder; infectious disease;  
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
XX myasthenia gravis.  
XX  
XX Homo sapiens.  
XX  
XX  
XX W0200192523-A2.  
XX  
XX 06-DEC-2001.  
XX  
XX 29-MAY-2001; 2001WO-US010836.  
XX  
XX 30-MAY-2000; 2000US-0206132P.  
XX 29-AUG-2000; 2000US-0228716P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shimkets RA, Leach MD;  
XX  
XX WPI; 2002-106308/14.  
XX N-PSDB; ABN23666.  
XX  
XX Novel human polypeptides and polynucleotides useful for diagnosing,  
XX preventing and treating cardiovascular disease, neurodegenerative,  
XX hyperproliferative disorders and autoimmune disorders.  
XX  
XX Disclosure; SEQ ID NO 15810; 1037pp; English.  
XX  
XX The present invention describes substantially purified human proteins  
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
XX in the specification)). ABN15762 to ABN27252 encode the human ORFX  
XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
XX treating or preventing a pathology associated with an ORFX-associated  
XX disorder in humans, and in the manufacture of a medicament for treating a  
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
XX sequences can be used in gene therapy. ORFX sequences can be used in the  
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
XX psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
XX osteoarthritis, neurodegenerative disorders, disorders related to organ  
XX transplantation, cardiovascular diseases, diabetes mellitus, systemic  
XX lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
XX storage disease, various immune deficiencies and disorders, infectious  
XX diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
XX arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
XX disease and autoimmune inflammatory eye disease. ORFX proteins are also  
XX useful for treating burns, incisions, ulcers, for treating osteoporosis,  
XX bone degenerative disorders, or periodontal disease, and for gut  
XX protection or regeneration and treatment of lung or liver fibrosis,  
XX reperfusion injury in various tissues and conditions resulting from  
XX systemic cytokine damage. N.B. The sequence data for this patent did not  
XX form part of the printed specification, but was obtained in electronic  
XX format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 67 AA;

Query Match

55.6%; Score 5; DB 5; Length 67;

Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
Db |||||  
47 FLRNP 51

## RESULT 35

AAU20901  
ID AAU20901 standard; protein; 68 AA.  
XX AC AAU20901;

XX 17-DEC-2001 (first entry)  
XX  
XX Human novel foetal antigen, SEQ ID NO 1145.

XX Human; foetal tissue antigen; antiinflammatory; neuroprotective;  
XX immunomodulator; cardiovascular; cytostatic; nephrothropic;  
XX cardiovascular; autoimmune disease; rheumatoid arthritis;  
XX hyperproliferative disorder; breast neoplasm; cancer;  
XX cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
XX cerebral ischaemia; angiogenesis; nervous system disorder;  
XX Alzheimer's disease; infection; ocular disorder; corneal infection;  
XX wound healing; epithelial cell proliferation; food additive.

Homo sapiens.

OS  
XX W0200155312-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US001321.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

24-FEB-2000; 2000US-0184664P.

02-MAR-2000; 2000US-0186350P.

16-MAR-2000; 2000US-0189874P.

17-MAR-2000; 2000US-0190076P.

18-APR-2000; 2000US-0198123P.

19-MAY-2000; 2000US-0205515P.

07-JUN-2000; 2000US-0209467P.

28-JUN-2000; 2000US-0214886P.

30-JUN-2000; 2000US-0215135P.

07-JUL-2000; 2000US-0216847P.

11-JUL-2000; 2000US-0217487P.

14-JUL-2000; 2000US-0217496P.

26-JUL-2000; 2000US-0220963P.

26-JUL-2000; 2000US-0220964P.

14-AUG-2000; 2000US-0224518P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225213P.

14-AUG-2000; 2000US-0225214P.

14-AUG-2000; 2000US-0225266P.

14-AUG-2000; 2000US-0225267P.

14-AUG-2000; 2000US-0225268P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225447P.

14-AUG-2000; 2000US-0225757P.

14-AUG-2000; 2000US-0225758P.

14-AUG-2000; 2000US-0225759P.

18-AUG-2000; 2000US-0226279P.

22-AUG-2000; 2000US-0226681P.

22-AUG-2000; 2000US-0226686P.

22-AUG-2000; 2000US-0227182P.

23-AUG-2000; 2000US-0227009P.

30-AUG-2000; 2000US-0228924P.

01-SEP-2000; 2000US-0229287P.

01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 25-SEP-2000; 2000US-0234988P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246533P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-488782/53.  
 DR N-PSDB; AAS33721.  
 XX  
 PT New polynucleotides and polypeptides for diagnosing, treating, preventing  
 PT or prognosing e.g. diseases or disorders of the nervous, musculoskeletal,  
 PT excretory, gastrointestinal, reproductive, and respiratory systems.  
 XX  
 PS Claim 11; SEQ ID NO 1145; 642pp; English.

CC The invention relates to novel nucleic acids encoding novel human foetal  
 CC antigens. The nucleic acids and proteins are used to prevent, treat (e.g.  
 CC by gene therapy) or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. The antibodies to the antigens can also be used  
 CC in alleviating symptoms associated with the disorders and in diagnostic  
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.  
 CC The polypeptides can also be used to aid wound healing and epithelial  
 CC cell proliferation, to prevent skin aging due to sunburn, to maintain  
 CC organs before transplantation, for supporting cell culture of primary  
 CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can  
 CC also be used as a food additive or preservative to increase or decrease  
 CC storage capabilities, fat content, lipid, protein, carbohydrate,  
 CC vitamins, minerals, cofactors and other nutritional components. Numerous  
 CC examples of diseases and disorders treated by the nucleic acids and  
 CC proteins are given in the specification. The present sequence represents  
 CC a foetal antigen of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained

Query Match 55.6%; Score 5; DB 4; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ESWFL 6  
 |||||  
 DB 18 ESWFL 22

RESULT 36  
AAM92279  
ID AAM92279 standard; protein; 69 AA.  
XX  
XX  
AC AAM92279;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
XX  
DE Human digestive system antigen SEQ ID NO: 1628.  
XX  
XX  
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;  
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
KW digestive system disorder; Meckel's diverticulum.  
XX  
OS Homo sapiens.  
XX  
PN W0200155314-A2.  
XX  
XX  
PD 02-AUG-2001.  
XX  
XX  
XX 17-JAN-2001; 2001WO-US001324.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239355P.  
PR 13-OCT-2000; 2000US-0239357P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.

```
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX N-PSDB; AAK80052.
XX
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
XX Claim 11; SEQ ID NO 1628; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a digestive system antigen of
CC the invention
XX
XX SQ Sequence 69 AA;
    Query Match          55.6%; Score 5; DB 4; Length 69;
    Best Local Similarity 100.0%; Pred. No. 2.1e+02;
    Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      3 SWFLR 7
Db      |||||
        19 SWFLR 23
RESULT 37
AAG19216
ID AAG19216 standard; protein; 73 AA.
XX
XX AAG19216;
AC
XX
XX DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 20933.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144003P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
```

PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144622P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147182P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149358P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155119P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 23-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.

PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161040P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 55.6%; Score 5; DB 3; Length 73;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
Db 44 FLRNP 48

## RESULT 38

AAG32128  
ID AAG32128 standard; protein; 73 AA.

XX AC AAG32128;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 38701.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131144P.  
PR 30-APR-1999; 99US-0131240P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140595P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142220P.  
PR 12-JUL-1999; 99US-0142777P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149923P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158023P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159325P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.



```

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160747P.
PR 21-OCT-1999; 99US-0160757P.
PR 21-OCT-1999; 99US-0160758P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 55.6%; Score 5; DB 3; Length 73;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
DB 44 FLRNP 48

RESULT 39
AAM15429
ID AAM15429 standard; protein; 79 AA.
XX AC AAM15429;
XX DT 12-OCT-2001 (first entry)
XX DE Peptide #1863 encoded by probe for measuring cervical gene expression.
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX OS cervical cancer.
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US0000670.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX Claim 27; SEQ ID NO 20255; 487bp; English.
XX The present invention relates to human single exon nucleic acid probes

```

```

CC (SENIP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENIPs are derived from human Hela cells. The SENIPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 79 AA;
SQ

Query Match 55.6%; Score 5; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7
DB 32 SWFLR 36

RESULT 40
ABB34441
ID ABB34441 standard; peptide; 79 AA.
XX AC ABB34441;
XX DT 04-FEB-2002 (first entry)
XX DE Peptide #1947 encoded by human foetal liver single exon probe.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US0000669.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human foetal liver.
XX Claim 27; SEQ ID NO 27076; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human foetal liver. The
XX present sequence is a peptide encoded by a single exon nucleic acid probe
XX of the invention. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 79 AA;
SQ

Query Match 55.6%; Score 5; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 SWFLR 7

Db 32 SWFLR 36

Search completed: August 31, 2006, 10:46:53  
Job time : 116.75 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:40:05 ; Search time 17.25 Seconds

(without alignments)  
50.200 Million cell updates/sec

Title: DENGUE\_SEROTYPE4

Perfect score: 9

Sequence: 1 veswflrnp 9

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 1

Total number of hits satisfying chosen parameters: 283392

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database :

PIR\_80.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	66.7	94	2 T29563	hypothetical prote
2	5	55.6	102	2 A12711	hypothetical prote
3	5	55.6	102	2 G97493	hypothetical prote
4	5	55.6	102	2 S19733	hypothetical prote
5	5	55.6	111	2 B82830	hypothetical prote
6	5	55.6	128	2 C84457	MHC sex-limited pr
7	5	55.6	130	2 C64097	hypothetical prote
8	5	55.6	132	2 F64097	hypothetical prote
9	5	55.6	156	2 B36905	fumarate reductase
10	5	55.6	187	2 AC1870	conserved hypochet
11	5	55.6	190	2 T37168	hypothetical prote
12	5	55.6	206	2 D95227	probable tetR-fam
13	5	55.6	206	2 H98091	hypothetical prote
14	5	55.6	208	2 T33341	hypothetical prote
15	5	55.6	217	2 I67411	hypothetical prote
16	5	55.6	217	2 A87990	somatotropin - rhe
17	5	55.6	219	2 C69439	protein W05H12.1 l
18	5	55.6	222	2 AE2003	sugar fermentation
19	5	55.6	230	2 B75280	hypothetical prote
20	5	55.6	236	2 E83879	probable phenylace
21	5	55.6	239	2 D84004	hypothetical prote
22	5	55.6	244	2 D96707	hypothetical prote
23	5	55.6	249	2 F83477	probable zinc fing
24	5	55.6	257	2 S65958	hypothetical prote
25	5	55.6	265	2 C45392	manuJ protein - Par
26	5	55.6	265	2 D36861	orf3 protein - por
27	5	55.6	265	2 D36861	orf3 protein - lel
28	5	55.6	282	2 AF2179	hypothetical prote
29	5	55.6	308	2 T29142	hypothetical prote

30	5	55.6	313	2 A65140	gtuNR operon regu
31	5	55.6	321	2 G86010	regulator of gluco
32	5	55.6	324	2 S74363	chlorophyll syntha
33	5	55.6	328	2 H64554	heat shock protein
34	5	55.6	331	2 AB0995	gluconate utilizat
35	5	55.6	331	2 G91164	regulator of gluco
36	5	55.6	338	2 T10038	hypothetical prote
37	5	55.6	340	2 D91154	hypothetical prote
38	5	55.6	340	2 A86000	hypothetical prote
39	5	55.6	340	2 D65129	hypothetical prote
40	5	55.6	343	2 H96783	hypothetical prote
41	5	55.6	358	2 E84467	hypothetical prote
42	5	55.6	376	2 A71175	probable dehydroge
43	5	55.6	385	2 E71238	hypothetical prote
44	5	55.6	388	2 AE2641	aspartate aminotra
45	5	55.6	391	2 G97423	probable aspartate
46	5	55.6	395	2 T39856	probable nadh-depe
47	5	55.6	397	2 D70512	hypothetical prote
48	5	55.6	403	2 S57945	probable translati
49	5	55.6	427	2 S17148	alpha-thrombin rec
50	5	55.6	427	2 S11367	UI snRNP 70K prote
51	5	55.6	432	2 A43448	thrombin receptor
52	5	55.6	440	2 F72038	RNA polymerase sig
53	5	55.6	440	2 A86587	RNA polymerase sig
54	5	55.6	443	2 T39497	hypothetical prote
55	5	55.6	447	2 AF2295	hypothetical prote
56	5	55.6	484	2 JC5779	4-carboxy-2-hydrox
57	5	55.6	485	2 JW0056	2-hydroxyuconic s
58	5	55.6	486	2 S10772	2-hydroxyuconic s
59	5	55.6	486	2 E42902	2-hydroxyuconic s
60	5	55.6	504	2 JC7613	cytochrome P450 2S
61	5	55.6	505	2 T31272	4-carboxy-2-hydrox
62	5	55.6	513	2 T05735	cytochrome P450 71
63	5	55.6	519	2 T45764	hypothetical prote
64	5	55.6	520	2 H69125	hypothetical prote
65	5	55.6	525	2 AD2022	hypothetical prote
66	5	55.6	526	1 PSXRB	outer capsid prote
67	5	55.6	526	2 S18768	outer capsid prote
68	5	55.6	557	2 S21733	FACC protein - hum
69	5	55.6	561	2 T05294	amidophosphoribos
70	5	55.6	575	2 AB1417	ABC transporter (A
71	5	55.6	575	2 AB1793	ABC transporter (A
72	5	55.6	584	2 G86713	hypothetical prote
73	5	55.6	591	2 I49656	Fanconi anemia gro
74	5	55.6	597	2 E85090	probable transposo
75	5	55.6	598	2 S66954	probable membrane
76	5	55.6	599	2 S67084	probable membrane
77	5	55.6	599	2 B72368	conserved hypochet
78	5	55.6	610	2 G86407	hypothetical prote
79	5	55.6	634	2 D70331	leucine-tRNA ligas
80	5	55.6	662	2 S42826	probable ATPase -
81	5	55.6	663	2 C82415	ATP-dependent RNA
82	5	55.6	683	2 A39784	phycobilisome anch
83	5	55.6	732	2 A43315	ETS domain protein
84	5	55.6	761	2 A46193	88K B-26-specific
85	5	55.6	802	2 T45642	FtsH metalloprotei
86	5	55.6	896	2 S76064	hypothetical prote
87	5	55.6	896	2 S59990	phycobilisome anch
88	5	55.6	911	2 S70911	transferrin-bindin
89	5	55.6	912	2 S70901	transferrin-bindin
90	5	55.6	912	2 C64107	transferrin-bindin
91	5	55.6	914	2 S70906	transferrin-bindin
92	5	55.6	915	2 E86514	CT131 homolog (imp
93	5	55.6	961	2 T03467	NADH dehydrogenase
94	5	55.6	971	2 T00268	hypothetical prote
95	5	55.6	995	2 T50267	probable family 31
96	5	55.6	1019	2 T30148	hypothetical prote
97	5	55.6	1080	2 A35088	phycobilisome link
98	5	55.6	1111	2 A59000	water protein (imp
99	5	55.6	1124	2 B84742	probable receptor-
100	5	55.6	1132	2 AD1809	phycobilisome core
101	5	55.6	1138	2 G71554	probable transmemb
102	5	55.6	1142	2 D72108	ct131 homolog-(pro

103 conserved hypothet 5 55.6 1142 2 F81562  
104 hypothetical prote 5 55.6 1395 2 T00068  
105 sex-limited protei 5 55.6 1735 2 S54784  
106 sex-limited protei 5 55.6 1736 2 A23176  
107 complement C4a pre 5 55.6 1738 1 C4HU  
108 myr 6, unconventio 5 55.6 1744 1 A24558  
109 sodium channel pro 5 55.6 1846 2 A59289  
110 hypothetical prote 5 55.6 2005 2 B25019  
111 hypothetical prote 5 55.6 2149 2 T47655  
112 hypothetical prote 5 55.6 2329 2 T28125  
113 delta-(L-alpha-ani 5 55.6 3071 2 T45584  
114 neuropeptide Antho 5 55.6 3770 2 A40889  
115 hypothetical prote 4 44.4 4 2 A35779  
116 hypothetical prote 4 44.4 23 2 E64634  
117 hypothetical prote 4 44.4 26 2 S14036  
118 hypothetical prote 4 44.4 26 2 S14037  
119 hypothetical prote 4 44.4 26 2 S13989  
120 hypothetical prote 4 44.4 28 2 C97078  
121 hypothetical prote 4 44.4 45 2 B70226  
122 rhodopsin fortytwo 4 44.4 51 4 I38158  
123 hypothetical prote 4 44.4 55 2 E82528  
124 hypothetical prote 4 44.4 58 2 A86702  
125 set1B protein - Sh 4 44.4 61 2 S54166  
126 hypothetical prote 4 44.4 61 2 T02733  
127 hypothetical prote 4 44.4 61 2 F96005  
128 hypothetical prote 4 44.4 61 2 AG2239  
129 hypothetical prote 4 44.4 63 2 T15583  
130 hypothetical prote 4 44.4 63 2 AB0097  
131 gene 12 protein - 4 44.4 64 2 S98141  
132 hypothetical prote 4 44.4 65 2 G90502  
133 lantibiotic cytoly 4 44.4 65 2 E83706  
134 M protein precurs 4 44.4 67 2 T16372  
135 hypothetical prote 4 44.4 68 2 S60797  
136 hypothetical prote 4 44.4 70 2 T18008  
137 genome polyprotein 4 44.4 71 2 PC1300  
138 genome polyprotein 4 44.4 71 2 PC1302  
139 hypothetical prote 4 44.4 71 2 T07190  
140 hypothetical prote 4 44.4 71 2 B64479  
141 hypothetical prote 4 44.4 72 2 T32629  
142 hypothetical prote 4 44.4 72 2 AE1875  
143 ATP synthase F0, C 4 44.4 74 2 F87294  
144 ATP synthase C cha 4 44.4 74 2 T17834  
145 ATP synthase chain 4 44.4 75 2 AE2664  
146 H+-transporting tw 4 44.4 75 2 D97446  
147 hypothetical prote 4 44.4 75 2 AC3445  
148 hypothetical prote 4 44.4 75 2 T16742  
149 relB protein - Esc 4 44.4 78 2 T17804  
150

ALIGNMENTS

RESULT 1  
T29563  
hypothetical protein T12E12.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
submitted to the EMBL Data Library, June 1996  
R:Bradshaw, H.; Stellyes, L.  
A:Description: The sequence of C. elegans cosmid T12E12.  
A:Reference number: Z20641  
A:Accession: T29563  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-94 <BRA>  
A:Cross-references: UNIPARC:UPI00001641F8; EMBL:U61944; PIDN:AAB03122.1; GSPDB:GN00022;  
A:Experimental source: strain Bristol N2; clone T12E12  
C:Genetics:  
A:Gene: CBSP.T12E12.5  
A:Map position: 4  
A:Introns: 63/3

Query Match 66.7%; Score 6; DB 2; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 SWFLRN 8  
DB 89 SWFLRN 94  
|||||  
RESULT 2  
A12711  
hypothetical protein Atul098 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: A12711  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;  
Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: A12711  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-102 <KUR>  
A:Cross-references: UNIPROT:Q8UGE0; UNIPARC:UPI00000D1A39; GB:AE008688; PIDN:AAL42111.1;  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atul098  
A:Map position: circular chromosome  
Query Match 55.6%; Score 5; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 SWFLR 7  
DB 8 SWFLR 12  
|||||

RESULT 3  
G97493  
hypothetical protein AGR\_C\_2032 [imported] - Agrobacterium tumefaciens (strain C58, Cerec  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: G97493  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, D.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: G97493  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-102 <KUR>  
A:Cross-references: UNIPROT:Q8UGE0; UNIPARC:UPI00000D1A39; GB:AE007869; PIDN:AAK86904.1;  
C:Genetics:  
A:Gene: AGR\_C\_2032  
A:Map position: circular chromosome  
Query Match 55.6%; Score 5; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 SWFLR 7  
DB 8 SWFLR 12  
|||||

RESULT 4

S19733  
 Hypothetical protein 2 - *Thiobacillus versutus* (fragment)  
 C;Species: *Thiobacillus versutus*  
 C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
 C;Accession: S19733  
 R;Ubbink, M.; van Kleef, M.A.G.; Kleinjan, D.J.; Hoitink, C.W.G.; Huitema, F.; Beintema, Eur. J. Biochem. 202, 1003-1012, 1991  
 A;Title: Cloning, sequencing and expression studies of the genes encoding amicyanin and A;Reference number: S19730; MUID:92111471; PMID:1765062  
 A;Accession: S19733  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-102 <UB>  
 A;Cross-references: UNIPROT:Q56464; UNIPARC:UPI00001704ED; GB:M58001; NID:g154632; PIDN:  
 Query Match 55.6%; Score 5; DB 2; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRNP 9  
 Db 37 FLRNP 41  
 RESULT 5  
 B82830  
 Hypothetical protein XF0247 [imported] - *Xylella fastidiosa* (strain 9a5c)  
 C;Species: *Xylella fastidiosa*  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C;Accession: B82830  
 R;anonymus, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
 A;Reference number: A82515; MUID:20365717; PMID:10910347  
 A;Note: for a complete list of authors see reference number A59328 below  
 A;Accession: B82830  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-111 <SIM>  
 A;Cross-references: UNIPROT:Q9PGQ1; UNIPARC:UPI00000C2346; GB:AE003878; GB:AE003849; NID  
 A;Experimental source: strain 9a5c  
 R;Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
 as-Neto, E.; Docena, C.; El-Dorfi, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
 Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 M.; Tshukato, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 A;Reference number: A59328  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: XF0247  
 Query Match 55.6%; Score 5; DB 2; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRNP 9  
 Db 69 FLRNP 73  
 RESULT 6  
 I69024  
 MHC sex-limited protein - mouse (fragment)  
 C;Species: *Mus musculus* (house mouse)  
 C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
 C;Accession: I69024

R;Nonaka, M.; Nakayama, K.; Yeul, Y.D.; Shimizu, A.; Takahashi, M.  
 Immunol. Rev. 87, 81-99, 1985  
 A;Title: Molecular cloning and characterization of complementary and genomic DNA clones  
 A;Reference number: I54567; MUID:86031969; PMID:2997024  
 A;Accession: I69024  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-128 <RES>  
 A;Cross-references: UNIPARC:UPI000011DE32; GB:M12974; NID:g199272; PIDN:AAA39560.1; PID:  
 C;Genetics:  
 A;Introns: 22/3; 86/3  
 C;Superfamily: alpha-2-macroglobulin  
 Query Match 55.6%; Score 5; DB 2; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRNP 9  
 Db 57 FLRNP 61  
 RESULT 7  
 C84457  
 Hypothetical protein Ar2g04410 [imported] - *Arabidopsis thaliana*  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C;Accession: C84457  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: C84457  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-130 <STO>  
 A;Cross-references: UNIPROT:Q9SJC8; UNIPARC:UPI000009FB78; GB:AE002093; NID:g4587609; PI  
 C;Genetics:  
 A;Gene: At2g04410  
 A;Map position: 2  
 Query Match 55.6%; Score 5; DB 2; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 FLRNP 9  
 Db 101 FLRNP 105  
 RESULT 8  
 F64097  
 fumarate reductase 15K protein homolog - *Haemophilus influenzae* (strain Rd KW20)  
 C;Species: *Haemophilus influenzae*  
 C;Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
 C;Accession: F64097  
 R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A  
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J  
 , D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
 Science 269, 496-512, 1995  
 A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J  
 A;Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.  
 A;Reference number: A64000; MUID:95350630; PMID:7542800  
 A;Accession: F64097  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-132 <TIG>  
 A;Cross-references: UNIPROT:P44892; UNIPARC:UPI000012ABE9; GB:U32765; GB:L42023; NID:g15  
 C;Genetics:  
 A;Start codon: GTG  
 C;Superfamily: fumarate reductase, subunit C (membrane anchor)

Query Match 55.6%; Score 5; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
|||||  
Db 65 FLRNP 69

RESULT 9  
B36905  
conserved hypothetical protein ylxS - *Bacillus subtilis*  
C:Species: *Bacillus subtilis*  
C/Date: 07-Apr-1994 #sequence revision 18-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: B36905; E65882; S31930  
R/Shazand, K.; Tucker, J.; Grunberg-Manago, M.; Rabinowitz, J.C.; Leighton, T.  
J. Bacteriol. 175, 2880-2887, 1993  
A/Title: Similar organization of the nusA-infB operon in *Bacillus subtilis* and *Escherichia coli*  
A/Reference number: A36905; MUID:93259931; PMID:8491709  
A/Accession: B36905  
A/Molecule type: DNA  
A/Residues: 1-156 <SHA>  
A/Cross-references: UNIPROT:P32726; UNIPARC:UPI0000060401; EMBL:Z18631; NID:g49314; PIDN:A36905  
A/Note: sequence extracted from NCBI backbone (NCBI:131760, NCBI:131765)  
R/Kunst, F.; Ogaawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berti, C.; Bron, S.; Broutillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chedoke, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997  
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.; Koester, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Magalhães, J.; M. Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrina, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A/Reference number: A69580; MUID:98044033; PMID:9384377  
A/Accession: B69882  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-156 <KUN>  
A/Cross-references: UNIPARC:UPI0000060401; GB:Z99112; GB:AL009126; NID:g26333902; PIDN:CA36905  
A/Experimental source: strain 168  
C/Genetics:  
A/Genes: ylxS  
C:Superfamily: nus operon 15K protein

Query Match 55.6%; Score 5; DB 2; Length 156;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7  
|||||  
Db 35 SWFLR 39

RESULT 10  
AC1870  
hypothetical protein alr0508 [imported] - *Nostoc* sp. (strain PCC 7120)  
C/Species: *Nostoc* sp. PCC 7120  
A/Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C/Accession: AC1870  
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriquchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.; DNA Res. 8, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AC1870

Db 153 FLRNP 157

RESULT 13

H98091

hypothetical protein spr1762 [imported] - Streptococcus pneumoniae (strain R6)

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C;Accession: H98091

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; H

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A;Reference number: A97872; MUID:21429245; PMID:11544234

A;Accession: H98091

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-206 <KUR>

A;Cross-references: UNIPROT:Q97NS2; UNIPARC:UPI0000051A5E; GB:AE007317;

C;Genetics:

A;Gene: spr1762

Query Match 55.6%; Score 5; DB 2; Length 206;

Best Local Similarity 100.0%; Pred. No. 59;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRNP 9

Db 153 FLRNP 157

RESULT 14

T33341

hypothetical protein K07D4.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 31-Dec-2004

C;Accession: T33341

R;Henkhaus, J.; Wohldmann, P.

submitted to the EMBL Data Library, July 1998

A;Description: The sequence of C. elegans cosmid K07D4.

A;Reference number: Z21327

A;Accession: T33341

A;Status: preliminary;

A;Molecule type: DNA

A;Residues: 1-208 <HEN>

A;Cross-references: UNIPROT:O76574; UNIPARC:UPI000007CD7A; EMBL:AF077534; PIDN:AAC26289.

A;Experimental source: strain Bristol N2; clone K07D4

C;Genetics:

A;Gene: CESP:K07D4.5

A;Map position: 2

A;Introns: 25/3; 68/1; 127/1; 160/2

Query Match 55.6%; Score 5; DB 2; Length 208;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WFLRN 8

Db 160 WFLRN 164

RESULT 15

I67411

sonatotropin - rhesus macaque

N;Alternate names: growth hormone

C;Species: Macaca mulatta (rhesus macaque)

C;Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004

C;Accession: I67411

R;Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.

Endocrinology 133, 1744-1752, 1993

A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementar

A;Reference number: I53267; MUID:94008724; PMID:8404617

A;Accession: I67411

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-217 <RES>

A;Cross-references: UNIPROT:Q07370; UNIPARC:UPI000016C489; GB:L16555; NID:g2931116; PIDN:?

C;Superfamily: prolactin

Query Match 55.6%; Score 5; DB 2; Length 217;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRNP 9

Db 70 FLRNP 74

RESULT 16

A87990

protein W05H12.1 [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C;Accession: A87990

R;anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A;Reference number: A75000; MUID:99069613; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_eleg

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A;Accession: A87990

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-217 <STO>

A;Cross-references: UNIPROT:O62394; UNIPARC:UPI000017A565; GB:chr\_I; PIDN:CAB04920.1; PII

C;Genetics:

A;Gene: W05H12.1

A;Map position: 1

Query Match 55.6%; Score 5; DB 2; Length 217;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SWFLR 7

Db 188 SWFLR 192

RESULT 17

C69439

sugar fermentation stimulation protein (sfsa) homolog - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: C69439

R;Kienk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.A

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: C69439

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-219 <KLE>

A;Cross-references: UNIPROT:O28756; UNIPARC:UPI0000056CBB; GB:AE000997; GB:AE000782; NID:

Query Match 55.6%; Score 5; DB 2; Length 219;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRNP 9

```
Db      209 FLRNP 213

RESULT 18
AE2003
hypothetical protein alr1579 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AE2003
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2003
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-222 <KUR>
A:Cross-references: UNIPROT:Q8YMW6; UNIPARC:UPI00000CE135; GB:BA0000119; PIDN:BAB77945.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1579

Query Match      55.6%; Score 5; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 SWFLR 7
      |||||
Db      5 SWFLR 9

RESULT 19
B75280
probable phenylacetic acid degradation protein PaaB - Deinococcus radiodurans (strain R1
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: B75280
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.O.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75280
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <WHI>
A:Cross-references: UNIPROT:Q9RRV0; UNIPARC:UPI00000D3FA4; GB:AE002069; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2385
A:Map position: 1

Query Match      55.6%; Score 5; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VESWF 5
      |||||
Db      172 VESWF 176

RESULT 20
E83879
hypothetical protein BH1837 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: E83879
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
```

```
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E83879
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-236 <STO>
A:Cross-references: UNIPROT:Q9KBT7; UNIPARC:UPI00000C3CF8; GB:AP001513; GB:BA000004; NID:
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1837
```

```
Query Match      55.6%; Score 5; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 ESWFL 6
      |||||
Db      14 ESWFL 18

RESULT 21
D84004
hypothetical protein BH2836 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: D84004
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D84004
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-239 <STO>
A:Cross-references: UNIPROT:Q9K915; UNIPARC:UPI00000C4008; GB:AP001516; GB:BA000004; NID:
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2836
```

```
Query Match      55.6%; Score 5; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 WFLRN 8
      |||||
Db      227 WFLRN 231
```

```
RESULT 22
D96707
probable zinc finger protein T22E19.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D96707
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96707
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <STO>
A:Cross-references: UNIPROT:Q9C9H1; UNIPARC:UPI00000A50C3; GB:AE005173; NID:g6715717; PI1
```

```
Query Match      55.6%; Score 5; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VESWF 5
      |||||
Db      172 VESWF 176

RESULT 20
E83879
hypothetical protein BH1837 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: E83879
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
```



Query Match 55.6%; Score 5; DB 2; Length 244;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
|||||  
Db 126 FLRNP 130

RESULT 23  
F83477  
hypothetical protein PA1350 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: F83477  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: F83477  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-249 <STO>  
A:Cross-references: UNIPROT:Q913Z4; UNIPARC:UPI00000C532D; GB:AE004564; GB:AE004091; NID:10984043  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA1350

Query Match 55.6%; Score 5; DB 2; Length 249;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
|||||  
Db 183 FLRNP 187

RESULT 24  
S65958  
mauJ protein - Paracoccus denitrificans (fragment)  
C:Species: Paracoccus denitrificans  
C>Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S65958  
R:van der Pallen, C.J.N.M.; Slotboom, D.J.; Jongejan, L.; Reijnders, W.N.M.; Harms, N.; Durr, J. Biochem. 230, 860-871, 1995  
A:Title: Mutational analysis of mau genes involved in methylamine metabolism in Paracoccus denitrificans  
A:Reference number: S65958; MUID:95324575; PMID:7601147  
A:Accession: S65958  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-257 <VAN>  
A:Cross-references: UNIPROT:P22566; UNIPARC:UPI000016FD2A; EMBL:U15028; NID:9595839; PMID:7601147  
C:Genetics:  
A:Gene: mauJ

Query Match 55.6%; Score 5; DB 2; Length 257;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
|||||  
Db 7 FLRNP 11

RESULT 25  
C45392  
orf3 protein - porcine reproductive and respiratory syndrome virus  
C:Species: porcine reproductive and respiratory syndrome virus, PRRSV  
C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 05-Oct-2004  
C:Accession: C45392  
R:Conzelmann, K.K.; Visser, N.; Van Woensel, P.; Thiel, H.J.

Query Match 55.6%; Score 5; DB 2; Length 244;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
|||||  
Db 126 FLRNP 130

RESULT 23  
F83477  
hypothetical protein PA1350 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: F83477  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: F83477  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-249 <STO>  
A:Cross-references: UNIPROT:Q913Z4; UNIPARC:UPI00000C532D; GB:AE004564; GB:AE004091; NID:10984043  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA1350

Query Match 55.6%; Score 5; DB 2; Length 249;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
|||||  
Db 183 FLRNP 187

RESULT 24  
S65958  
mauJ protein - Paracoccus denitrificans (fragment)  
C:Species: Paracoccus denitrificans  
C>Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S65958  
R:van der Pallen, C.J.N.M.; Slotboom, D.J.; Jongejan, L.; Reijnders, W.N.M.; Harms, N.; Durr, J. Biochem. 230, 860-871, 1995  
A:Title: Mutational analysis of mau genes involved in methylamine metabolism in Paracoccus denitrificans  
A:Reference number: S65958; MUID:95324575; PMID:7601147  
A:Accession: S65958  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-257 <VAN>  
A:Cross-references: UNIPROT:P22566; UNIPARC:UPI000016FD2A; EMBL:U15028; NID:9595839; PMID:7601147  
C:Genetics:  
A:Gene: mauJ

Query Match 55.6%; Score 5; DB 2; Length 257;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
|||||  
Db 7 FLRNP 11

RESULT 25  
C45392  
orf3 protein - porcine reproductive and respiratory syndrome virus  
C:Species: porcine reproductive and respiratory syndrome virus, PRRSV  
C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 05-Oct-2004  
C:Accession: C45392  
R:Conzelmann, K.K.; Visser, N.; Van Woensel, P.; Thiel, H.J.

Query Match 55.6%; Score 5; DB 2; Length 265;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7  
|||||  
Db 196 SWFLR 200

RESULT 26  
D36861  
orf3 protein - Lelystad virus  
C:Species: Lelystad virus  
C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 05-Oct-2004  
C:Accession: D36861  
R:Meulenbergh, J.J.; Hulst, M.M.; de Meijer, E.J.; Moonen, P.L.; den Besten, A.; de Kluyver, E. J. Virol. 192, 62-72, 1993  
A:Title: Lelystad virus, the causative agent of porcine epidemic abortion and respiratory disease in pigs  
A:Reference number: A44281; MUID:93297139; PMID:8517032  
A:Contents: annotation  
A>Note: no sequence in this paper  
C:Superfamily: uncharacterized conserved protein

Query Match 55.6%; Score 5; DB 2; Length 265;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7  
|||||  
Db 196 SWFLR 200

RESULT 27  
S42125  
hypothetical protein 3 - Mycoplasma capricolum  
C:Species: Mycoplasma capricolum  
C>Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S42125  
R:Miyata, M.; Sano, K.I.; Okada, R.; Fukumura, T.  
Nucleic Acids Res. 21, 4816-4823, 1993  
A:Title: Mapping of replication initiation site in Mycoplasma capricolum genome by two-dimensional DNA sequencing  
A:Reference number: S42116; MUID:94051609; PMID:8233831  
A:Accession: S42125  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-279 <MIY>  
A:Cross-references: UNIPROT:P43045; UNIPARC:UPI000013C014; EMBL:D14982; NID:9416237; PMID:8233831  
C:Genetics:  
A:Genetic code: SGC3

Query Match 55.6%; Score 5; DB 2; Length 279;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      2  ESWFL 6
      |||||
Db      87  ESWFL 91

RESULT 28
AF2179
hypothetical protein all2989 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF2179
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Itiguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2179
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-282 <KUR>
A:Cross-references: UNIPROT:Q8YSU3; UNIPARC:UPI00000CE61F; GB:BA000019; PIDN:BAB74688.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all2989

Query Match      55.6%; Score 5; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5  FLRNP 9
      |||||
Db      247  FLRNP 251

RESULT 29
T29142
hypothetical protein K11C4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29142
R:Pauley, A.; Gattung, S.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid K11C4.
A:Reference number: Z20577
A:Accession: T29142
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-308 <PAU>
A:Cross-references: UNIPROT:Q94277; UNIPARC:UPI000007D331; EMBL:U64854; PIDN:AAB18314.1;
A:Experimental source: strain Bristol N2; clone K11C4
C:Genetics:
A:Gene: CESP:K11C4.1
A:Map position: 5
A:Introns: 56/3; 81/1; 212/2; 265/2
C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match      55.6%; Score 5; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  VESWF 5
      |||||
Db      209  VESWF 213

RESULT 30
A65140
gtnr operon regulator - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: A65140
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

```

```

A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97428617; PMID:9278503
A:Accession: A65140
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-313 <BLAT>
A:Cross-references: UNIPROT:P48860; UNIPARC:UPI00001681BC; GB:AE000420; GB:U000096; NID:g
A:Experimental source: strain K-12, substrain MGL655
C:Genetics:
A:Gene: gnrR
C:Superfamily: gnrR protein

Query Match      55.6%; Score 5; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5  FLRNP 9
      |||||
Db      25  FLRNP 29

RESULT 31
G86010
regulator of gluconate (gnt) operon [imported] - Escherichia coli (strain O157:H7, subst
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: G86010
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551.
A:Accession: G86010
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <STO>
A:Cross-references: UNIPROT:Q8X6V8; UNIPARC:UPI0000165962; GB:AE005174; NID:g12518088; P
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: gnrR
C:Superfamily: gnrR protein

Query Match      55.6%; Score 5; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5  FLRNP 9
      |||||
Db      15  FLRNP 19

RESULT 32
S74363
chlorophyll synthase chain 33K - Synechocystis sp. (strain PCC 6803)
N:Alternate names: hypothetical protein slr0056
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S74363
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O.K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74363
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-324 <KAN>
A:Cross-references: UNIPROT:Q55145; UNIPARC:UPI00000347CC; EMBL:D64001; GB:AB001339; NID:
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

```

## C;Genetics:

A;Gene: G4  
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0279

Query Match 55.6%; Score 5; DB 2; Length 324;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRNP 9  
|||||  
Db 289 FLRNP 293

RESULT 33  
H64554  
heat shock protein B - Helicobacter pylori (strain 26695)  
C;Species: Helicobacter pylori  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 05-Oct-2004  
C;Accession: H64554  
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: H64554  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-328 <TOM>  
A;Cross-references: UNIPROT:O25057; UNIPARC:UPI000000D2FD9; GB:AE000546; GB:AE000511; NID  
C;Superfamily: lipid A biosynthesis (KDO)2-(lauroyl)-lipid IVA acyltransferase

## Query Match

Best Local Similarity 55.6%; Score 5; DB 2; Length 328;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SWFLR 7

Db 49 SWFLR 53  
|||||

## RESULT 34

AB0995  
gluconate utilization operon repressor [imported] - Salmonella enterica subsp. enterica  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AB0995  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J. S.; Moule, S.; O'Gaora, P. Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AB0995  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-331 <PAR>  
A;Cross-references: UNIPARC:UPI000000A579D; GB:AL513382; PTDN:CAD08086.1; PID:g16505065;  
C;Genetics:  
A;Gene: STV4268  
C;Superfamily: gntR protein

## Query Match

Best Local Similarity 55.6%; Score 5; DB 2; Length 331;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRNP 9

Db 25 FLRNP 29  
|||||

## RESULT 35

G91164  
regulator of gluconate operon [imported] - Escherichia coli (strain O157:H7, substrain R1)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: G91164  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: G91164  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-331 <HAY>  
A;Cross-references: UNIPROT:Q8X6V8; UNIPARC:UPI000000E6E3; GB:BA0000007; PTDN:BA037710.1;  
A;Experimental source: strain O157:H7, substrain R1MD 0509952  
C;Genetics:  
A;Gene: ECs4287  
C;Superfamily: gntR protein

## Query Match

Best Local Similarity 55.6%; Score 5; DB 2; Length 331;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRNP 9

Db 25 FLRNP 29  
|||||

## RESULT 36

T10038  
hypothetical protein MLCB628.19c - Mycobacterium leprae  
C;Species: Mycobacterium leprae  
C;Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: T10038  
R;Biglmeier, K.; Honore, N.; Woods, S.A.; Caudron, B.; Cole, S.T. Mol. Microbiol. 7, 197-206, 1993  
A;Title: Use of an ordered cosmid library to deduce the genomic organization of Mycobacter  
A;Reference number: Z16917; MUID:93188700; PMID:8446027  
A;Accession: T10038  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-338 <BIG>  
A;Cross-references: UNIPROT:O33090; UNIPARC:UPI000000AF0A0; EMBL:Y14967; NID:g2370268; PFI  
C;Genetics:  
A;Note: MLCB628.19c

## Query Match

Best Local Similarity 55.6%; Score 5; DB 2; Length 338;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FLRNP 9

Db 64 FLRNP 68  
|||||

## RESULT 37

D91154  
hypothetical protein ECs4204 [imported] - Escherichia coli (strain O157:H7, substrain R1)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: D91154  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: D91154  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-340 <HAY>

A:Cross-references: UNIPROT:Q8X873; UNIPARC:UPI00000D0989; GB:BA000007; PIDN:BAB37627.1;  
 A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetics:  
 A:Gene: ECs4204  
 C:Superfamily: alpha/beta hydrolase, YheT type

Query Match 55.6%; Score 5; DB 2; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7  
 |||||  
 Db 132 SWFLR 136

## RESULT 38

hypothetical protein yheT [imported] - Escherichia coli (strain O157:H7, substrain EDL93)  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: A86000

R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: A86000  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-340 <STO>  
 A:Cross-references: UNIPROT:Q8X873; UNIPARC:UPI00000D0989; GB:AE005174; NID:g12517975; E  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: yheT  
 C:Superfamily: alpha/beta hydrolase, YheT type

Query Match 55.6%; Score 5; DB 2; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7  
 |||||  
 Db 132 SWFLR 136

## RESULT 39

D65129  
 hypothetical 38.5 kD protein in kifb-prkb intergenic region - Escherichia coli (strain K  
 C:Species: Escherichia coli  
 C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
 C:Accession: D65129

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
 .A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: D65129  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
 A:Residues: 1-340 <BLAT>  
 A:Cross-references: UNIPROT:P45524; UNIPARC:UPI000013B219; GB:AE000411; GB:U00096; NID:Q  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: yheT  
 C:Superfamily: alpha/beta hydrolase, YheT type

Query Match 55.6%; Score 5; DB 2; Length 340;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7  
 |||||  
 Db 132 SWFLR 136

## RESULT 40

H96783

hypothetical protein FlB16.12 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: H96783

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96783

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-343 <STO>

A:Cross-references: UNIPROT:Q9FWS3; UNIPARC:UPI000009E921; GB:AE005173; NID:g10120449; P

C:Genetics:

A:Gene: FlB16.12

A:Map position: 1

Query Match 55.6%; Score 5; DB 2; Length 343;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9  
 |||||  
 Db 185 FLRNP 189

Search completed: August 31, 2006, 10:48:01  
 Job time : 21.25 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 10:29:54 ; Search time 139.25 Seconds  
(without alignments)  
59.786 Million cell updates/sec

Title: DENGUE\_SEROTYPE4

Perfect score: 9

Sequence: 1 veswflrnp 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2849581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database :

Uniprot 7.2.2\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	66.7	216	2	Q8VPM7_MICCC
2	6	66.7	238	2	Q8Q599_PSEPK
3	6	66.7	272	2	Q72TK1_LEPIC
4	6	66.7	272	2	Q8F1Q1_LEPIN
5	6	66.7	275	2	Q746R7_GEOSL
6	6	66.7	365	2	Q2IQ56_9DELT
7	6	66.7	407	2	Q7XMI7_ORYSA
8	6	66.7	451	2	Q2R4T2_ORYSA
9	6	66.7	1035	2	Q3AB12_CARHZ
10	5	55.6	55	2	Q4XEM2_PLACH
11	5	55.6	70	2	Q54AN7_DICDI
12	5	55.6	73	2	Q93I28_ARATH
13	5	55.6	73	2	Q7UQW7_RHOBA
14	5	55.6	78	1	ICP47_HSV2S
15	5	55.6	78	2	Q5YON0_9ALPH
16	5	55.6	86	2	Q3P9Q0_PARDE
17	5	55.6	90	2	Q6LAQ9_AMICA
18	5	55.6	102	2	Q8UGB0_AGRYS
19	5	55.6	106	2	Q3X8M6_METFL
20	5	55.6	106	2	Q8HHB1_BRAFA
21	5	55.6	111	2	Q9PGQ1_XYLFA
22	5	55.6	120	2	Q4JJK6_ANOGA
23	5	55.6	120	2	Q4JUL5_ANOGA
24	5	55.6	120	2	Q4JUL6_ANOGA
25	5	55.6	120	2	Q4JJL7_ANOGA
26	5	55.6	120	2	Q4JJM4_ANOGA
27	5	55.6	120	2	Q4JUN3_ANOGA
28	5	55.6	120	2	Q4JJP9_ANOGA
29	5	55.6	122	2	Q34W18_9GAMM
30	5	55.6	122	2	Q4SDC8_9BURK
31	5	55.6	122	2	Q4LWW9_9BURK

32	5	55.6	122	2	Q3U251_MOUSE
33	5	55.6	127	2	Q9DIU9_MOUSE
34	5	55.6	128	2	Q32IF4_9NEOB
35	5	55.6	130	2	Q9SJC8_ARATH
36	5	55.6	132	1	FRDC_HAEIN
37	5	55.6	134	2	Q3QPM1_9RHOB
38	5	55.6	136	2	Q4QM67_HAEI8
39	5	55.6	137	1	RUVX_BUCBP
40	5	55.6	138	2	Q6SHD5_9BACT
41	5	55.6	139	2	Q3HEU6_TRIER
42	5	55.6	142	2	Q3LX02_9HIV1
43	5	55.6	142	2	Q3LX18_9HIV1
44	5	55.6	142	2	Q58GL1_9HIV1
45	5	55.6	143	2	Q4TBN4_TETNG
46	5	55.6	149	2	Q48HM6_PSEL4
47	5	55.6	156	1	YLXS_BACSU
48	5	55.6	157	2	Q388Q6_9TRYP
49	5	55.6	157	2	Q3XZ67_ENTFC
50	5	55.6	157	2	Q62UZO_BACLD
51	5	55.6	159	2	Q7VY38_BORPE
52	5	55.6	159	2	Q7WB12_BORPA
53	5	55.6	159	2	Q7WMH7_BORBR
54	5	55.6	165	2	Q4MQ36_BACCE
55	5	55.6	166	2	Q65U15_BACLD
56	5	55.6	167	2	Q427F6_DESHA
57	5	55.6	172	2	Q58DG8_BOVIN
58	5	55.6	172	2	Q43RJ4_SOLUS
59	5	55.6	173	2	Q5SJN7_THET8
60	5	55.6	175	2	Q72K11_THET2
61	5	55.6	177	2	Q37XB6_SPHAR
62	5	55.6	180	2	Q4CT10_TRYCR
63	5	55.6	180	2	Q4DQM4_TRYCR
64	5	55.6	181	2	Q7X9R7_HORVD
65	5	55.6	186	2	Q6ZEV6_SYNV3
66	5	55.6	187	2	Q8YZF2_ANASP
67	5	55.6	189	2	Q64EK1_9ARCH
68	5	55.6	189	2	Q4SV42_TETNG
69	5	55.6	190	2	Q9S1N8_STRCO
70	5	55.6	190	2	Q9DSU2_9VIRU
71	5	55.6	191	2	Q81VS3_HUMAN
72	5	55.6	192	2	Q977R8_9CREN
73	5	55.6	194	2	Q3GC31_9FIRM
74	5	55.6	194	2	Q3KXK0_9PROT
75	5	55.6	195	2	Q7PSEL_ANOGA
76	5	55.6	197	2	Q31LM5_SYNP7
77	5	55.6	197	2	Q3QSO0_9RHOB
78	5	55.6	197	2	Q5N099_SYNP6
79	5	55.6	198	2	Q3GC97_9FIRM
80	5	55.6	198	2	Q6MB92_PARUW
81	5	55.6	199	2	Q2NCP9_9SPHN
82	5	55.6	206	2	Q8CYA2_STRR6
83	5	55.6	207	2	Q4B3C0_9BURK
84	5	55.6	207	2	Q4B3C0_9BURK
85	5	55.6	207	2	Q4B3C0_9BURK
86	5	55.6	207	2	Q4B3C0_9BURK
87	5	55.6	207	2	Q4B3C0_9BURK
88	5	55.6	207	2	Q4B3C0_9BURK
89	5	55.6	207	2	Q4B3C0_9BURK
90	5	55.6	207	2	Q4B3C0_9BURK
91	5	55.6	207	2	Q4B3C0_9BURK
92	5	55.6	207	2	Q4B3C0_9BURK
93	5	55.6	207	2	Q4B3C0_9BURK
94	5	55.6	207	2	Q4B3C0_9BURK
95	5	55.6	207	2	Q4B3C0_9BURK
96	5	55.6	207	2	Q4B3C0_9BURK
97	5	55.6	207	2	Q4B3C0_9BURK
98	5	55.6	207	2	Q4B3C0_9BURK
99	5	55.6	207	2	Q4B3C0_9BURK
100	5	55.6	207	2	Q4B3C0_9BURK
101	5	55.6	207	2	Q4B3C0_9BURK
102	5	55.6	207	2	Q4B3C0_9BURK
103	5	55.6	207	2	Q4B3C0_9BURK
104	5	55.6	207	2	Q4B3C0_9BURK

Q3U251	mus musculus
Q9DIU9	mus musculus
Q32IF4	phasmahyla
Q9SJC8	arabidopsis
P4492	haemophilus
Q3QPM1	silicibacte
Q4QM67	haemophilus
Q89A50	buchnera ap
Q6SHD5	uncultured
Q3HEU6	trichodesmi
Q3LX02	human immun
Q3LX18	human immun
Q58GL1	human immun
Q4TBN4	tetradon n
Q48HM6	pseudomonas
P32726	bacillus su
Q388Q6	trypanosoma
Q3XZ67	enterococcu
Q62UZO	bacillus li
Q7VY38	bordetella
Q7WB12	bordetella
Q7WMH7	bordetella
Q4MQ36	bacillus ce
Q65U15	bacillus li
Q427F6	desulfitoba
Q58DG8	bos taurus
Q43RJ4	solibacter
Q5SJN7	thermus the
Q72K11	thermus the
Q37XB6	novosphingo
Q4CT10	trypanosoma
Q4DQM4	trypanosoma
Q7X9R7	hordeum vul
Q6ZEV6	synchocyst
Q8YZF2	anabaena sp
Q64EK1	uncultured
Q4SV42	tetradon n
Q9S1N8	strepomyce
Q4C776	crocosphae
Q9S1N8	strepomyce
Q9DSU2	diadromus p
Q81VS3	homo sapien
Q977R8	uncultured
Q3GC31	syntrophomo
Q3KXK0	magnetococc
Q7PSEL	anopheles g
Q31LM5	synchococc
Q3QSO0	silicibacte
Q5N099	synchococc
Q3GC97	syntrophomo
Q6MB92	parachlamyd
Q2NCP9	erythrobaet
Q4TEP6	tetradon n
Q5UPC5	mimivirus.
Q3XRM8	magnetococc
Q8CYA2	strepococc
Q97NS2	strepococc
Q4B3C0	polaromonas
Q4B3C0	polaromonas
Q2P521	xanthomonas
Q7QPH7	giardia lam
Q5H262	xanthomonas
Q8P6J1	xanthomonas
Q8PJ91	xanthomonas
Q41XA8	desulfitoba
Q07370	macaca mula
Q31315	alouatta ee
Q4VUI4	macaca asse
Q4VUI5	macaca asse
Q8MI74	callithrix
Q8WND9	ateles geof
Q28A60	hahella che
Q28756	archaeoglob
Q5JF11	homo sapien
Q94QQ3	inversidens



```

OX NCBI_TaxID=44275;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Fiocruz L1-130;
RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monceliro-Vitello C.B.,
RA Ho P.U., Haake D.A., Verjovsky-Almeida S., Hartskeerl R.A.,
RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorri H.,
RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A.,
RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
RA Jeronimo S.M.B., Junqueira-de-Azevedo J.L.M., Kimura E.T.,
RA Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
RA de Oliveira R.C., Pereira G.G., Reis M.S., Schrieffer A.,
RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
RT "Comparative genomics of two Leptospira interrogans serovars reveals
RT novel insights into physiology and pathogenesis.";
RL J. Bacteriol. 186:2164-2172(2004).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AE016823; AAS69427.1; -; Genomic_DNA.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR006087; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
KW Complete proteome.
SQ SEQUENCE 272 AA; 32759 MW; 49C5F6760BB406DF CRC64;

Query Match 66.7%; Score 6; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLRN 8
Db 201 SWFLRN 206
|||||

RESULT 4
ID Q8F1Q1_LEPIN PRELIMINARY; PRT; 272 AA.
AC Q8F1Q1;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-MAR-2006, entry version 13.
DE Sterol desaturase family protein.
GN OrderedLocusNames=LA3078; ORFNames=LA_3078;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AE010300; AAN50276.1; -; Genomic_DNA.
DR GenomeReviews; AE010300; LA3078.

```

```

DR BioCyc; LINT189518:LA3077-MONOMER; -.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006087; Sterol_desat.
DR InterPro; IPR006088; Sterol_desatur.
DR Pfam; PF01598; Sterol_desat; 1.
KW Complete proteome.
SQ SEQUENCE 272 AA; 32746 MW; E6C5F7330BB410B3 CRC64;

Query Match 66.7%; Score 6; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLRN 8
Db 201 SWFLRN 206
|||||

RESULT 5
ID Q746R7_GEOSL PRELIMINARY; PRT; 275 AA.
AC Q746R7;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Hypothetical protein.
GN OrderedLocusNames=GSU3451;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Winn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.E., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA Van Aken S.E., Lovley D.R., Fraser C.M.;
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments.";
RL Science 302:1967-1969(2003).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AB017180; AAR36841.1; -; Genomic_DNA.
DR TIGR; GSU3451; -.
DR BioCyc; GSU35554.GSU3451-MONOMER; -.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000379; Ser_estr.
KW Complete proteome; Hydrolase; Hypothetical protein.
SQ SEQUENCE 275 AA; 30210 MW; B5A972E003C99EEB CRC64;

Query Match 66.7%; Score 6; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VESWFL 6
Db 61 VESWFL 66
|||||

RESULT 6
ID Q2IQ56_9DELT PRELIMINARY; PRT; 365 AA.
AC Q2IQ56;
DT 07-MAR-2006, integrated into UniProtKB/TrEMBL.
DT 07-MAR-2006, sequence version 1.
DT 07-MAR-2006, entry version 1.

```



```
DE Acyltransferase 3.
GN ORNames=Adeh.1161;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacteriineae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute;
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter J.C., Glavina T.,
RA Hammon N., Istrani S., Pitluck S., Brettin T., Bruce D., Han C.,
RA Tapia R., Gilna P., Kiss H., Schmutz J., Larimer F., Land M.,
RA Kyrpides N., Anderson I., Sanford R.A., Ritalahti K.M., Thomas H.S.,
RA Kirby J.R., Zhulin I.B., Loeffler F.E., Richardson P.,
RA "Complete sequence of Anaeromyxobacter dehalogenans 2CP-C."
RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000251; ABC80935.1; -; Genomic DNA.
KW Acyltransferase; Transferase.
SQ SEQUENCE 365 AA; 39837 MW; DF5D43FDC9240D3 CRC64;
Query Match 66.7%; Score 6; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ESWFLR 7
DB 337 ESWFLR 342
RESULT 7
Q7XMI7 ORYSA
ID Q7XMI7 ORYSA PRELIMINARY; PRT; 407 AA.
AC Q7XMI7;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2004, sequence version 2.
DE 07-FEB-2006, entry version 11.
DE OSJNB0006N15.12 protein.
GN OSJNB0006N15.12;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RL "Sequence and analysis of rice chromosome 4.";
RT Nature 420:316-320(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AL607003; CAE04595.2; -; Genomic DNA.
DR HSSP; Q13231; 1GUU.
DR Gramene; Q7XMI7.
DR GO; GO:0004568; F:chitinase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006032; P:chitin catabolism; IEA.
```

```
DR InterPro; IPR011583; Chitinase II.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
SQ SEQUENCE 407 AA; 42455 MW; 871BC26EB7B856F8 CRC64;
Query Match 66.7%; Score 6; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SWFLRN 8
DB 288 SWFLRN 293
RESULT 8
Q2R4T2 ORYSA
ID Q2R4T2 ORYSA PRELIMINARY; PRT; 451 AA.
AC Q2R4T2;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DE 07-FEB-2006, entry version 2.
DE Glycosyl hydrolases family 18, putative.
GN ORNames=LOC_Os11g27400;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wing R.A., McCombie W.A., Ouyang S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DP000010; ABA93461.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
KW Hydrolase.
SQ SEQUENCE 451 AA; 49352 MW; B57A86A9CB81737A CRC64;
Query Match 66.7%; Score 6; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SWFLRN 8
DB 287 SWFLRN 292
RESULT 9
Q3AB12 CARHZ
ID Q3AB12 CARHZ PRELIMINARY; PRT; 1035 AA.
AC Q3AB12;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Molybdopterin oxidoreductase, molybdopterin-binding subunit.
GN OrderedLocustNames=CHY_1852;
OS Carboxydotherrus hydrogenoformans (strain Z-2901 / DSM 6008).
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae;
OC Carboxydotherrus.
OX NCBI_TaxID=246194;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16311624; DOI=10.1371/journal.pgen.0010065;
RA Wu M., Ren Q., Durkin A.S., Daugherty S.C., Brinkac L.M., Dodson R.J.,
RA Madupu R., Sullivan S.A., Kolonay J.F., Nelson W.C., Tallon L.J.,
RA Jones K.M., Sullivan S.A., Gonzalez J.M., Zhulin I.B., Robb F.T.,
RA Eisen J.A.;
RL "Life in hot carbon monoxide: the complete genome sequence of
RT Carboxydotherrus hydrogenoformans Z-2901.";
```



```

RL PLoS Genet. 1:563-574(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CP000141; AB15360.1; -; Genomic_DNA.
DR TIGR; CHY 1852; -.
DR GO; GO:0030151; F:molybdenum ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR KW Complete proteome.
DR SQ SEQUENCE 1035 AA; 115121 MW; AA24A61217A94CEC CRC64;

Query Match 66.7%; Score 6; DB 2; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VESWFL 6
Db 40 VESWFL 45

RESULT 10
Q4XEM2 PLACH PRELIMINARY; PRT; 55 AA.
AC Q4XEM2;
DT 05-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Hypothetical protein.
GN ORFNames=PC402496.00.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15637271; DOI=10.1126/science.1103717;
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Beriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.M.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R. III,
RA Kafatos F.C., Janse C.J., Barrell B.G., Turner C.M.R., Waters A.P.,
RA Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDJB whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAJ01006961; CAH84646.1; -; Genomic_DNA.
DR KW Hypothetical protein.
DR SQ SEQUENCE 55 AA; 7661 MW; A74199959FC3PBBE CRC64;

Query Match 55.6%; Score 5; DB 2; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
Db 32 FLRNP 36

RESULT 11
Q54AN7 D1CDI
ID Q54AN7 D1CDI PRELIMINARY; PRT; 70 AA.
AC Q54AN7;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.

```

```

DE Hypothetical protein.
GN ORFNames=DDB0215156;
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Suganb R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Farhorthou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Karbhornou P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsey R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulsegod H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT "The genome of the social amoeba Dictyostellium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDJB whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAF101000323; EAL60324.1; -; Genomic_DNA.
DR KW Hypothetical protein.
DR SQ SEQUENCE 70 AA; 8181 MW; 8770C0E45B8CDEDB CRC64;

Query Match 55.6%; Score 5; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7
Db 17 SWFLR 21

RESULT 12
Q93YZ8 ARATH PRELIMINARY; PRT; 73 AA.
AC Q93YZ8;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE At2G04410/T103.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DDJB databases.
RN [2]

```

```

RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.W., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
RA Toriumi M., Wallender E.K., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinzaki K., Davis R.W., Theologos A., Ecker J.R.;
RA Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RL
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY058877; RAL24264.1; -; mRNA.
DR EMBL; EF001029; AAN4783.1; -; mRNA.
SQ SEQUENCE 73 AA; 8505 MW; EEBAA685EE95287 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
Db 44 FLRNP 48

RESULT 13
Q7UQW7_RHOBA
ID Q7UQW7_RHOBA PRELIMINARY; PRT; 73 AA.
AC Q7UQW7;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Hypothetical protein.
GN OrderedLocustNames=RB6038;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RA "Complete genome sequence of the marine planctomycete Pirellula sp.
RA strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BX294143; CAD74578.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 73 AA; 8560 MW; 5A14B65F06460F2A CRC64;

Query Match 55.6%; Score 5; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ESWFL 6
Db 54 ESWFL 58

RESULT 14
ICP47_HSV2S
ID ICP47_HSV2S STANDARD; PRT; 78 AA.
AC F60504;
DT 01-MAR-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-MAR-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.

```

```

DE ICP47 protein (Infected cell protein 47) (Immediate-early protein
DE IE12) (Immediate-early-5) (Vmw12) (US12 protein).
GN Name=US12;
OS Herpes simplex virus type 2 (strain SA8) (Simian agent 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10316;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=Isolate B264;
RA Bigger J.E., Martin D.W.;
RL "Identification of an ICP47 homolog in Simian agent 8 (SA8).";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds specifically to transporters associated with
CC antigen processing (TAP), thereby blocking peptide-binding and
CC translocation by TAP as well as subsequent loading of peptides
CC onto MHC class I molecules in the endoplasmic reticulum. In
CC consequence, infected cells are masked for immune recognition by
CC cytotoxic T lymphocytes (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm (By similarity).
CC -!- DOMAIN: The N-terminal active domain blocks peptide binding to and
CC peptide transport by TAP (By similarity).
CC -!- SIMILARITY: Belongs to the herpesviruses US12 family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY387672; AAQ90018.1; -; Genomic DNA.
KW Early protein; Viral immunoevasion.
FT CHAIN 1 78 ICP47 protein.
FT REGION 3 36 /FTID=PRO_0000115814.
FT SEQUENCE 78 AA; 8572 MW; 2A3942BF70406125 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 78;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
Db 12 FLRNP 16

RESULT 15
Q5Y0N0_9ALPH
ID Q5Y0N0_9ALPH PRELIMINARY; PRT; 78 AA.
AC Q5Y0N0;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE Immediate early protein ICP47.
GN Name=US12;
OS Cercopithecine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10317;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15629785; DOI=10.1016/j.virol.2004.09.042;
RA Tyler S.D., Peters G.A., Severini A.;
RT "Complete genome sequence of cercopithecine herpesvirus 2 (SA8) and
RT comparison with other simplexviruses.";
RL Virology 331:429-440(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY714813; AAU88138.1; -; Genomic DNA.
SQ SEQUENCE 78 AA; 8572 MW; 2A3942BF70406125 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      5 FLRNP 9
DB      12 FLRNP 16

RESULT 16
Q3P9Q0 PARDE PRELIMINARY; PRT; 86 AA.
AC Q3P9Q0;
DT 25-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=PDNDRAFT.0875;
OS Paracoccus denitrificans PD1222.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Paracoccus.
OX NCBI_TaxID=318586;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PD1222;
RG US DOE Joint Genome Institute (JGI-PGP);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Paracoccus
RT denitrificans PD1222.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PD1222;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Paracoccus denitrificans
RT PD1222.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AAIT01000016; EMBL64359.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 86 AA; 9530 MW; 035D5536940B3EDD CRC64;

Query Match 55.6%; Score 5; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 FLRNP 9
DB      62 FLRNP 66

RESULT 17
Q6LAQ9 AMICA PRELIMINARY; PRT; 90 AA.
AC Q6LAQ9;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Amia calva (Bowfin).
OG Euteleostomi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Amiiformes; Amiidae; Amia.
OX NCBI_TaxID=7924;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Birstein V.J., Hanner R., Desalle R.;

```

```

RT "Phylogeny of the Acipenseriformes: cytogenetic and molecular
RT aspects.";
RL Environ. Biol. Fishes 48:127-155(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Birstein V.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -! SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -! SIMILARITY: Belongs to the cytochrome b family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; X95060; CAA64466.1; -; Genomic_DNA.
DR SMR; Q6LAQ9; 1-89.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom B_N; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 10241 MW; B564B94639920CB4 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 WFLRN 8
DB      42 WFLRN 46

RESULT 18
Q8UGEO AGRTS PRELIMINARY; PRT; 102 AA.
AC Q8UGEO; Q7CZ27;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Hypothetical protein Atu1098 (AGR_C2032p).
GN OrderedLocustNames=AGR_C2032, Atu1098;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Woo L.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Bovee D. Sr.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;

```

RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 RT C58 ";  
 RL Science 294:2317-2323(2001).  
 RN [2]

RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RP MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houdiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty J., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT Agrobacterium tumefaciens C58";  
 RL Science 294:2323-2328(2001).

CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; AE009073; RA142111.1; -; Genomic\_DNA.  
 DR EMBL; AE008039; AK86904.1; -; Genomic\_DNA.  
 DR PIR; A12711; A12711.  
 DR PIR; G97493; G97493.

KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 102 AA; 11833 MW; 1858C364A6B5D911 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLR 7  
 |||||  
 DB 8 SWFLR 12

RESULT 19

Q3X8M6 METFL  
 ID Q3X8M6\_METFL PRELIMINARY; PRT; 106 AA.  
 AC Q3X8M6;  
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.  
 DT 11-OCT-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 3.  
 DE Hypothetical protein.  
 GN ORFNames=MFLADRAFT 0811;  
 OS Methylobacillus flagellatus KT.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Methylophilales;  
 OC Methylophilaceae; Methylobacillus.  
 OX NCBI\_TaxID=265072;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=KT;  
 RG US DOE Joint Genome Institute (JGI-PGF);  
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,  
 RA Hammon N., Istrani S., Ptluck S., Richardson P.;  
 RT "Sequencing of the draft genome and assembly of Methylobacillus  
 RT flagellatus KT.";  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=KT;  
 RG US DOE Joint Genome Institute (JGI-ORNL);  
 RA Larimer P., Land M.;  
 RT "Annotation of the draft genome assembly of Methylobacillus  
 RT flagellatus KT.";  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=KT;  
 RG US DOE Joint Genome Institute (JGI-PGF);  
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,  
 RA Hammon N., Istrani S., Ptluck S., Richardson P.;  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

KW Hypothetical protein.  
 SQ SEQUENCE 106 AA; 11797 MW; 8AAECE90D0E03909 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ESFWL 6  
 |||||  
 DB 12 ESFWL 16

RESULT 20

Q8HHB1 BRAFA  
 ID Q8HHB1\_BRAFA PRELIMINARY; PRT; 106 AA.  
 AC Q8HHB1;  
 DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.  
 DT 01-MAR-2003, sequence version 1.  
 DT 07-FEB-2006, entry version 15.  
 DE Cytochrome b (Fragment).  
 GN Name=cytb;  
 GN Brachylophus fasciatus (Fiji banded iguana).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Iguania; Iguanidae; Brachylophus.  
 OX NCBI\_TaxID=46195;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Houlden B.A., Costello B.H.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase  
 CC complex (complex III or cytochrome b-c1 complex), which is a  
 CC respiratory chain that generates an electrochemical potential  
 CC coupled to ATP synthesis (By similarity).  
 CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).  
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,  
 CC cytochrome c1 and the Rieske protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the cytochrome b family.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----

DR EMBL; AF459043; AA015566.1; -; Genomic\_DNA.

DR SMR; Q8HHB1; 1-106  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005506; F:iron ion binding; IEA.  
 DR GO; GO:0046872; F:metal ion binding; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR005797; Cytb\_b6\_N.  
 DR Pfam; PF00033; Cytochrom B\_N; 1.  
 DR PROSITE; PS51002; CYTB\_NTER; 1.  
 DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;  
 KW Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 106  
 SQ SEQUENCE 106 AA; 11894 MW; 14F14C7CDEBB1490 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRN 8

```
Db 111111
47 WFLRN 51

RESULT 21
Q9PGQ1_XYLFA PRELIMINARY; PRT; 111 AA.
AC Q9PGQ1;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Hypothetical protein.
GN OrderedLocusNames=Xf0247;
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=985C;
RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barroes M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Honeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.P., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT The genome sequence of the plant pathogen Xylella fastidiosa.;
RL Nature 406:151-159(2000).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; AB003878; AAF83060.1; -; Genomic_DNA.
DR EPR; B82830; B82830.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 111 AA; 12524 MW; AFC5CB4D03672AD5 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
Db 69 FLRNP 73
111111

RESULT 22
Q4JJK6_ANOGA PRELIMINARY; PRT; 120 AA.
AC Q4JJK6;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.

Query Match 55.6%; Score 5; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
Db 69 FLRNP 73
111111

RESULT 23
Q4JUL5_ANOGA PRELIMINARY; PRT; 120 AA.
AC Q4JUL5;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Olfactory receptor 38 (Fragment).
GN Names=GPCR38;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S3, S5, S6, and S13;
RX PubMed=16076241; DOI=10.1371/journal.pbio.0030285;
RA Turner T.L., Hahn M.W., Nuzhdin S.V.;
RT "Genomic islands of speciation in Anopheles gambiae.";
RL PLOS Biol. 3:E285-E285(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; DQ080897; AAY89543.1; -; Genomic DNA.
DR EMBL; DQ080901; AAY89547.1; -; Genomic DNA.
DR EMBL; DQ080903; AAY89549.1; -; Genomic DNA.
DR EMBL; DQ080893; AAY89539.1; -; Genomic DNA.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13865 MW; 896B435407EA60BB CRC64;

Query Match 55.6%; Score 5; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRN 8
Db 99 WFLRN 103
111111

Query Match 55.6%; Score 5; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRN 8
Db 99 WFLRN 103
111111

RESULT 24
Q4JUL5_ANOGA PRELIMINARY; PRT; 120 AA.
AC Q4JUL5;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Olfactory receptor 38 (Fragment).
GN Names=GPCR38;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S3, S5, S6, and S13;
RX PubMed=16076241; DOI=10.1371/journal.pbio.0030285;
RA Turner T.L., Hahn M.W., Nuzhdin S.V.;
RT "Genomic islands of speciation in Anopheles gambiae.";
RL PLOS Biol. 3:E285-E285(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; DQ080897; AAY89543.1; -; Genomic DNA.
DR EMBL; DQ080901; AAY89547.1; -; Genomic DNA.
DR EMBL; DQ080903; AAY89549.1; -; Genomic DNA.
DR EMBL; DQ080893; AAY89539.1; -; Genomic DNA.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13865 MW; 896B435407EA60BB CRC64;

Query Match 55.6%; Score 5; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRN 8
Db 99 WFLRN 103
111111
```

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 WFLRN 8
Db 99 WFLRN 103

RESULT 24
Q4JUL6 ANOGA
ID Q4JUL6 ANOGA PRELIMINARY; PRT; 120 AA.
AC Q4JUL6;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Olfactory receptor 38 (Fragment).
GN Name=GPCR38;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S3, S5, and S13;
RX PubMed=16076241; DOI=10.1371/journal.pbio.0030285;
RA Turner T.L., Hahn M.W., Nuzhdin S.V.;
RT "Genomic islands of speciation in Anopheles gambiae.";
RL PLoS Biol. 3:E285-E285(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; DQ080896; AAY89542.1; -; Genomic DNA.
DR EMBL; DQ080900; AAY89546.1; -; Genomic DNA.
DR EMBL; DQ080892; AAY89538.1; -; Genomic DNA.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13890 MW; 896B435410EC60BB CRC64;

Query Match 55.6%; Score 5; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 WFLRN 8
Db 99 WFLRN 103

RESULT 25
Q4JUL7 ANOGA
ID Q4JUL7 ANOGA PRELIMINARY; PRT; 120 AA.
AC Q4JUL7;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Olfactory receptor 38 (Fragment).
GN Name=GPCR38;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S2;
RX PubMed=16076241; DOI=10.1371/journal.pbio.0030285;
RA Turner T.L., Hahn M.W., Nuzhdin S.V.;
RT "Genomic islands of speciation in Anopheles gambiae.";
RL PLoS Biol. 3:E285-E285(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; DQ080896; AAY89542.1; -; Genomic DNA.
DR EMBL; DQ080900; AAY89546.1; -; Genomic DNA.
DR EMBL; DQ080892; AAY89538.1; -; Genomic DNA.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13890 MW; 896B435410EC60BB CRC64;

Query Match 55.6%; Score 5; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 WFLRN 8
Db 99 WFLRN 103

RESULT 26
Q4JUN3 ANOGA
ID Q4JUN3 ANOGA PRELIMINARY; PRT; 120 AA.
AC Q4JUN3;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Olfactory receptor 38 (Fragment).
GN Name=GPCR38;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S10, S12, S2, S4, S7, S8, S9, and S1;
RX PubMed=16076241; DOI=10.1371/journal.pbio.0030285;
RA Turner T.L., Hahn M.W., Nuzhdin S.V.;
RT "Genomic islands of speciation in Anopheles gambiae.";
RL PLoS Biol. 3:E285-E285(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; DQ080888; AAY89534.1; -; Genomic DNA.
DR EMBL; DQ080889; AAY89535.1; -; Genomic DNA.
DR EMBL; DQ080890; AAY89536.1; -; Genomic DNA.
DR EMBL; DQ080894; AAY89540.1; -; Genomic DNA.
DR EMBL; DQ080898; AAY89544.1; -; Genomic DNA.
DR EMBL; DQ080899; AAY89545.1; -; Genomic DNA.
DR EMBL; DQ080905; AAY89551.1; -; Genomic DNA.
DR EMBL; DQ080907; AAY89553.1; -; Genomic DNA.
DR EMBL; DQ080908; AAY89554.1; -; Genomic DNA.
DR EMBL; DQ080886; AAY89532.1; -; Genomic DNA.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13864 MW; 896B43540A5C60BB CRC64;

Query Match 55.6%; Score 5; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 WFLRN 8
Db 99 WFLRN 103

RESULT 27
Q4JUN3 ANOGA
ID Q4JUN3 ANOGA PRELIMINARY; PRT; 120 AA.
AC Q4JUN3;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
```

```

DT 02-AUG-2005, sequence version 1.
DE 07-FEB-2006, entry version 4.
DE Olfactory receptor 38 (Fragment).
GN Name=GPROR38;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MS;
RX PubMed=16076241; DOI=10.1371/journal.pbio.0030285;
RA Turner T.L., Hahn M.W., Nuzhdin S.V.;
RT "Genomic islands of speciation in Anopheles gambiae.";
RL PLoS Biol. 3:E285-E285(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ080879; AAY89525.1; -; Genomic_DNA.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13873 MW; 382B1752B79CCCD1 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRN 8
DB 99 WFLRN 103

RESULT 28
ID Q4JJP9 ANOGA PRELIMINARY; PRT; 120 AA.
AC Q4JJP9;
DT 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DE 07-FEB-2006, entry version 5.
DE Olfactory receptor 38 (Fragment).
GN Name=GPROR38;
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M1, M10, M11, M15, M17, M2, M3, M5, M6, M7, M8, M9, S1, S12,
RC S6, and S9;
RX PubMed=16076241; DOI=10.1371/journal.pbio.0030285;
RA Turner T.L., Hahn M.W., Nuzhdin S.V.;
RT "Genomic islands of speciation in Anopheles gambiae.";
RL PLoS Biol. 3:E285-E285(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ080863; AAY89509.1; -; Genomic_DNA.
DR EMBL; DQ080864; AAY89510.1; -; Genomic_DNA.
DR EMBL; DQ080865; AAY89511.1; -; Genomic_DNA.
DR EMBL; DQ080866; AAY89512.1; -; Genomic_DNA.
DR EMBL; DQ080867; AAY89513.1; -; Genomic_DNA.
DR EMBL; DQ080868; AAY89514.1; -; Genomic_DNA.
DR EMBL; DQ080869; AAY89515.1; -; Genomic_DNA.
DR EMBL; DQ080870; AAY89516.1; -; Genomic_DNA.
DR EMBL; DQ080871; AAY89517.1; -; Genomic_DNA.
DR EMBL; DQ080872; AAY89518.1; -; Genomic_DNA.
DR EMBL; DQ080873; AAY89519.1; -; Genomic_DNA.

```

```

DR EMBL; DQ080874; AAY89520.1; -; Genomic_DNA.
DR EMBL; DQ080875; AAY89521.1; -; Genomic_DNA.
DR EMBL; DQ080876; AAY89522.1; -; Genomic_DNA.
DR EMBL; DQ080877; AAY89523.1; -; Genomic_DNA.
DR EMBL; DQ080878; AAY89524.1; -; Genomic_DNA.
DR EMBL; DQ080880; AAY89526.1; -; Genomic_DNA.
DR EMBL; DQ080881; AAY89527.1; -; Genomic_DNA.
DR EMBL; DQ080882; AAY89528.1; -; Genomic_DNA.
DR EMBL; DQ080883; AAY89529.1; -; Genomic_DNA.
DR EMBL; DQ080884; AAY89530.1; -; Genomic_DNA.
DR EMBL; DQ080885; AAY89531.1; -; Genomic_DNA.
DR EMBL; DQ080887; AAY89533.1; -; Genomic_DNA.
DR EMBL; DQ080891; AAY89537.1; -; Genomic_DNA.
DR EMBL; DQ080902; AAY89548.1; -; Genomic_DNA.
DR EMBL; DQ080909; AAY89555.1; -; Genomic_DNA.
DR EMBL; DQ080862; AAY89508.1; -; Genomic_DNA.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13891 MW; 896B43541D5A60BB CRC64;

Query Match 55.6%; Score 5; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRN 8
DB 99 WFLRN 103

RESULT 29
ID Q34W18 9GAMM PRELIMINARY; PRT; 122 AA.
AC Q34W18;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Hypothetical protein.
GN ORFNames=MigDRAFT_0463;
OS Alkalilimnicola ehrlichei MLHE-1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;
OC Ectothiorhodospiraceae; Alkalilimnicola.
OX NCBI_TaxID=187272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MLHE-1;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler J.C., Glavina T.,
RA Hammon N., Istrani S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Alkalilimnicola
RT ehrlichei MLHE-1.";
RL Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MLHE-1;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome of Alkalilimnicola ehrlichei MLHE-1.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAK01000021; EAP33377.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 122 AA; 14286 MW; CC968883708AFD1B CRC64;

Query Match 55.6%; Score 5; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;

```







RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
 RA Ambesi-Impombato A., Apweiler K.W., Beraano T., Bono H., Chalk A.M.,  
 RA Bansal M., Baxter L., Beisel K.W., Christoffels A., Clutterbuck D.R.,  
 RA Chiu K.P., Choudhary V., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Huminicki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Matsuda H., Matsumura S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,  
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Rest B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Seesa L., Sheng Y.,  
 RA Shibata Y., Shinada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,  
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
 RA Yamashiki H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
 RA Wahlstedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawaohima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,  
 RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RA "The transcriptional landscape of the mammalian genome.";  
 RT Science 309:1559-1563(2005).  
 RL Science 309:1559-1563(2005).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD;  
 RX PubMed=16141073; DOI=10.1126/science.1112009;  
 RG RIKEN Genome Exploration Research Group, and Genome Science Group  
 RG (Genome Network Core Team) and the FANTOM Consortium;  
 RT "Antisense Transcription in the Mammalian Transcriptome.";  
 RL Science 309:1564-1566(2005).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaio I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Glasi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reid J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaio I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [7]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitesunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [8]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NOD;  
 RX STRAIN=NOD;  
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,  
 RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,  
 RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,  
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,  
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 CC EMBL: AK155491; BAE33291.1; -; mRNA.  
 DR Hypothetical protein  
 KW SEQUENCE 122 AA; 14369 MW; DD960A1D968FCB12 CRC64;  
 SQ  
 Query Match 55.6%; Score 5; DB 2; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VESWF 5  
 DB 101 VESWF 105

RESULT 33  
 Q9DIU9 MOUSE  
 ID Q9DIU9\_MOUSE PRELIMINARY; PRT; 127 AA.  
 AC Q9DIU9;  
 DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.  
 DT 01-JUN-2001, sequence version 1.  
 DT 07-FEB-2006, entry version 19.  
 DE 0 day neonate kidney cDNA, RIKEN full-length enriched library,  
 DE clone: D630002J15 product: hypothetical protein, full insert sequence.  
 GN Name=D630002J15R1K;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridea; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Methods Enzymol. 303:19-44(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX PubMed=16141072; DOI=10.1126/science.1112014;  
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,  
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,  
 RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,  
 RA Ambesi-Impombato A., Anweiler K., Aturaliya R.N., Bailey T.L.,  
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,  
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,  
 RA Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,  
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,  
 RA Guscinich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
 RA Hill D., Hummel K., Iacono M., Ikeo K., Iwama A., Ishikawa T.,  
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,  
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,  
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,  
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
 RA Rest B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
 RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
 RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,  
 RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,  
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,  
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,  
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,  
 RA Kawashina T., Kohjima M., Kondo S., Konno H., Nakano K., Ninomiya N.,  
 RA Nishio T., Okada M., Plesky C., Shibata K., Shiraki T., Suzuki S.,  
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,  
 RA Hayashizaki Y.;  
 RT "The transcriptional landscape of the mammalian genome.";  
 RL Science 309:1559-1563(2005).  
 RN [3]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX PubMed=16141073; DOI=10.1126/science.1112009;  
 RG RIKEN Genome Exploration Research Group, and Genome Science Group  
 RT "Antisense Transcription in the Mammalian Transcriptome.";  
 RL Science 309:1564-1566(2005).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Osato N., Saito K., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Kagnaga A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kigawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischnmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzei J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).

```

RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishimi T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Inotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AK021320; BAB32372.1; -: mRNA.
DR Ensembl: ENSMUSG00000029829; Mus musculus.
DR MGI: MGI:1924772; D6300020315Rik.
KW Hypothetical protein.
SQ SEQUENCE 127 AA; 13618 MW; 82155594064C9C1 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
DB 18 FLRNP 22

RESULT 34
Q32IF4_9NEOB PRELIMINARY; PRT; 128 AA.
AC Q32IF4;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Cytochrome b (Fragment).
GN Names=cytb;
OS Phasmanyula guttata (spotted leaf frog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;
OC Phyllomedusinae; Phasmanyula.
OX NCBI_TaxID=318407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Faivovich J., Haddad C.F.B., Garcia P.C.A., Frost D.R., Campbell J.A.,
RA Wheeler W.C.;
RT "Systematic Review of the frog family Hylidae, with special reference
RT to the Hylinae: Phylogenetic analysis and taxonomic revision.";
RL Bull. Am. Mus. Nat. Hist. 294:1-240(2005).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential

```

```

CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AY843961; AAX53355.1; -: Genomic DNA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR GO: GO:0005506; F:iron ion binding; IEA.
DR GO: GO:0046872; F:oxidoreductase activity; IEA.
DR GO: GO:0016491; F:metal ion binding; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR005797; Cytb_b6_N.
DR Pfam: PF00033; Cytochrom B_N; 1.
DR PROSITE: P851002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 128
SQ SEQUENCE 128 AA; 14431 MW; 4B3A934082C5C0D2 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRN 8
DB 71 WFLRN 75

RESULT 35
Q9SJC8_ARATH PRELIMINARY; PRT; 130 AA.
AC Q9SJC8;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Hypothetical protein At2g04410.
GN OrderedLocustNames=At2g04410;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.D., Ronning C.M., Benito M.-I.,
RA Carreira A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AC006951; AAD25837.1; -: Genomic DNA.
DR PIR: C84457; C84457.
DR TAIR: At2g04410; -.
KW Hypothetical protein.
SQ SEQUENCE 130 AA; 15034 MW; 7820AD483192B5D3 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;

```

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
Db 101 FLRNP 105

RESULT 36
FRDC_HAEIN STANDARD; PRT; 132 AA.
AC P44892;
DT 01-NOV-1995, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1995, sequence version 1.
DT 07-MAR-2006, entry version 38.
DE Fumarate reductase subunit C.
GN Name=frdC; OrderedLocNames=HI0833;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Keiley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Spriggs D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: Seems to be involved in the anchoring of the catalytic
CC components of the fumarate reductase complex to the cytoplasmic
CC membrane (By similarity).
CC -!- SUBUNIT: Part of an enzyme complex containing four subunits: a
CC flavoprotein (frdA), an iron-sulfur protein (frdB), and two
CC hydrophobic anchor proteins (frdC and frdD) (By similarity).
CC -!- SUBCELLULAR LOCATION: Bacterial cell inner membrane; multi-pass
CC membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the frdC family.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; L42023; AAC22491.1; -; Genomic_DNA.
CC HSP; P03805; I1K6.
CC DR PIR; P64097; F64097.
CC DR GenomeReviews; L42023_GR; HI0833.
CC DR TIGR; HI0833; -.
CC DR BioCyc; HINF71421:HI0833-MONOMER; -.
CC DR HAMAP; MF 00708; -; 1.
CC DR InterPro; IPR003510; Fumarate_red_C.
CC DR Pfam; PF02300; Fumarate_red_C; 1.
CC DR PIRSF; PIRSF00180; FrdC; 1.
CC DR ProDom; PD015900; Fumarate_red_C; 1.
CC KW Complete proteome; Inner membrane; Membrane; Transmembrane.
FT CHAIN 1 132 Fumarate reductase subunit C.
FT FTID=PRO_0000196530.
FT TRANSMEM 33 55 Potential.
FT TRANSMEM 70 92 Potential.
FT TRANSMEM 113 131 Potential.
SQ SEQUENCE 132 AA; 15265 MW; 03740E833BB58C4A CRC64;

Query Match 55.6%; Score 5; DB 1; Length 132;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9

```

```

Db 65 FLRNP 69

RESULT 37
Q3QPM1_9RHOB PRELIMINARY; PRT; 134 AA.
AC Q3QPM1;
DT 25-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 25-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Thioesterase superfamily.
GN ORFNames=RoseDRAFT_0603;
OS Silicibacter sp. TM1040.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxID=292414;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TM1040;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Silicibacter sp.
RT TM1040.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TM1040;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Silicibacter sp. TM1040.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TM1040;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pittluck S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; AAF02000009; EAN55569.1; -; Genomic_DNA.
CC GO; GO:0003824; F:catalytic activity; IEA.
CC DR InterPro; IPR006683; Thioestr_supf.
CC DR Pfam; PF03061; 4HBT; 1.
CC DR SEQUENCE 134 AA; 15390 MW; EA9AA9798B22DF94 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VESWF 5
Db 32 VESWF 36

RESULT 38
Q4QM67_HAEI8 PRELIMINARY; PRT; 136 AA.
ID Q4QM67_HAEI8
AC Q4QM67;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Fumarate reductase subunit C.
GN Name=frdC; OrderedLocNames=NTHI0999;
OS Haemophilus influenzae (strain 86-028NP).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```

```

OC Pasteurellaceae; Haemophilus.
RN NCBI_TaxID=281310;
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,
RA Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,
RA Munson R.S. Jr.;
RT "Genomic sequence of an otitis media isolate of nontypeable
RT Haemophilus influenzae: comparative study with H. influenzae serotype
RL d, strain KW20.";
RL J. Bacteriol. 187:4627-4636 (2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: CP000057; AAX87880.1; -; Genomic_DNA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR003510; Fumarate_red_C.
DR Pfam: PF02300; Fumarate_red_C; 1.
DR PIRSF: PIRSF000180; FrdC; 1.
DR ProDom: PD015900; Fumarate_red_C; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 136 AA; 15647 MW; 1973C0C155D9E4AE CRC64;

Query Match 55.6%; Score 5; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FLRNP 9
Db 69 FLRNP 73

RESULT 39
RUVX_BUCBP STANDARD; PRT; 137 AA.
ID Q89A50;
AC Q89A50;
DT 16-JUN-2003, integrated into UniProtKB/Swiss-Prot.
DT 16-JUN-2003, sequence version 1.
DT 07-MAR-2006, entry version 20.
DE Putative Holliday junction resolvase (EC 3.1.-.-).
GN OrderedLocuNames=bbp492;
OS Buchnera aphidicola subsp. Baizongia pistaciae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22426901; PubMed=12522265; DOI=10.1073/pnas.0235981100;
RA van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Iatorre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586 (2003).
CC -!- FUNCTION: Could be a nuclease that resolves Holliday junction
CC intermediates in genetic recombination.
CC -!- SUBCELLULAR LOCATION: Cytoplasm (Potential).
CC -!- SIMILARITY: Belongs to the yggf HJR family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AE016826; AA027197.1; -; Genomic_DNA.
DR GenomeReviews; AE016826 GR; bbp492.
DR BioCyc; BAPH224915:BBP492-MONOMER; -.
DR HAMAP; MF_00651; -.
DR InterPro: IPR005227; HJR_Yggf.
DR InterPro: IPR012337; RNaseH_fold.
DR InterPro: IPR006641; YggfC.
DR Pfam; PF03652; UPF0081; 1.
DR SMART; SM00732; YggfC; 1.

```

```

DR TIGRFAMS; TIGR00250; HJR_Yggf; 1.
KW Complete proteome; DNA damage; DNA recombination; DNA repair;
KW Hydrolase; Nuclease.
FT CHAIN 1 137 Putative Holliday junction resolvase.
FT /FTId=PRO_0000172036;
SQ SEQUENCE 137 AA; 15606 MW; 54484DBB43E1E05F CRC64;

Query Match 55.6%; Score 5; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ESWFL 6
Db 129 ESWFL 133

RESULT 40
Q6SHD5_9BACT PRELIMINARY; PRT; 138 AA.
ID Q6SHD5;
AC Q6SHD5;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Hypothetical protein.
GN ORFNames=EBAC750-02H05.3;
OS uncultured bacterium 440.
OC Bacteria; environmental samples.
OX NCBI_TaxID=257390;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA DeLong E.F.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AY458637; AAR37685.1; -; Genomic_DNA.
DR InterPro: IPR009562; DUF1178.
DR Pfam; PF06676; DUF1178; 1.
DR PIRSF: PIRSF032131; UCP032131; 1.
KW Hypothetical protein.
SQ SEQUENCE 138 AA; 15923 MW; 0EBA7FF1D348B9B7 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ESWFL 6
Db 15 ESWFL 19

Search completed: August 31, 2006, 10:39:37
Job time : 148.25 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 11:33:43 ; Search time 110.25 Seconds  
(without alignments)  
37.324 Million cell updates/sec

Title: DENGUE\_SEROTYPE1

Perfect score: 55

Sequence: 1 vetflrhp 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_8.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003as.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*
- 9: Geneseq2005s.\*
- 10: Geneseq2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	98.2	39	9	Adw12582 M1-40/DEN
2	54	98.2	48	9	Adw12588
3	48	87.3	9	9	Adw12595
4	48	87.3	21	9	Adw12594 M32-40/DE
5	48	87.3	32	9	Adw12593
6	48	87.3	39	9	Adw12576 M1-40/DEN
7	48	87.3	40	5	Aae17432 Dengue (D
8	48	87.3	40	5	Aae17433 (95-114)E
9	48	87.3	48	5	Aae17433
10	48	87.3	167	8	Adn37497 Dengue vi
11	48	87.3	171	8	Adn37493
12	48	87.3	171	8	Adn37496 Dengue vi
13	48	87.3	635	2	Aaw75410 Fusion pr
14	48	87.3	675	8	Adn37628 Dengue vi
15	48	87.3	675	8	Adn37518
16	48	87.3	675	8	Adn37612
17	48	87.3	675	8	Adn37626 Dengue vi
18	48	87.3	677	2	Aaw75411 Fusion pr
19	48	87.3	677	8	Adn37613 Dengue vi
20	48	87.3	681	8	Adn37603
21	48	87.3	681	8	Adn37517 Dengue vi
22	48	87.3	685	6	Abp57874 Plasmid p
23	48	87.3	685	6	Abp57876 Plasmid p

97 42 76.4 1131 4 ABG11655 Novel hum  
98 42 76.4 1232 7 ADF70474 Orphan re  
99 41 74.5 27 8 ADN11192 Peptide m  
100 41 74.5 27 8 ADN11216 Peptide m

## ALIGNMENTS

RESULT 1  
ADW12582  
ID ADW12582 standard; peptide; 39 AA.  
XX  
AC ADW12582;  
XX  
DT 24-MAR-2005 (first entry)  
XX  
DE M1-40/DEN-2 (F36) mutant protein.  
XX  
KW Gene therapy; protein purification; virucide; cytostatic; vaccine;  
KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
KW DEN; dengue; mutant; mutein.  
XX  
OS Dengue virus.  
XX  
PN US2004266987-A1.  
XX  
PD 30-DEC-2004.  
XX  
PF 30-JUN-2003; 2003US-00608029.  
XX  
PR 30-JUN-2003; 2003US-00608029.  
XX  
PA (INSP ) INST PASTEUR.  
XX  
PI Despres P, Catteau A;  
XX  
PI WPI; 2005-047647/05.  
XX  
DR New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
PT as a vaccine for preventing or treating pathological conditions from non-  
PT specific febrile illnesses to severe hemorrhagic manifestations or  
PT encephalitic syndromes.  
XX  
PS Example 1; SEQ ID NO 29; 30pp; English.  
XX  
CC The present invention relates to an isolated and purified ApoptoM  
CC peptide. The invention is useful as a vaccine for the prevention and  
CC treatment of pathological conditions from non-specific febrile illnesses  
CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
CC pathological conditions are linked to Flavivirus infection or cancers.  
CC The invention is also useful in gene therapy. The present sequence is a  
CC M1-40/DEN (dengue)-2 (F36) mutant protein.  
XX  
SQ Sequence 39 AA;

Query Match 98.2%; Score 54; DB 9; Length 39;  
Best Local Similarity 88.9%; Pred. No. 0.025;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
:|||||  
Db 31 IETWFLRHP 39

RESULT 2  
ADW12588  
ID ADW12588 standard; protein; 48 AA.  
XX  
AC ADW12588;  
XX  
XX  
DT 24-MAR-2005 (first entry)  
XX

DE p (95-114) EGFP (M1-M40)DEN-2 (136F) plasmid DNA encoded protein #3.  
XX Gene therapy; protein purification; virucide; cytostatic; vaccine;  
KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
KW DEN; dengue; EGFP; enhanced green fluorescent protein.  
XX  
OS Dengue virus.  
OS Chimeric.  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "Encoded by GGC"  
FT  
FT Misc-difference 4 /note= "Encoded by GAC"  
FT  
FT Misc-difference 13.44 /note= "Encoded by GTTTC"  
FT  
XX US2004266987-A1.  
XX  
XX 30-DEC-2004.  
PD  
XX  
PF 30-JUN-2003; 2003US-00608029.  
XX  
PR 30-JUN-2003; 2003US-00608029.  
XX  
PA (INSP ) INST PASTEUR.  
XX  
PI Despres P, Catteau A;  
XX  
PI WPI; 2005-047647/05.  
XX  
DR N-PSDB; ADW12589.  
XX  
PT New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
PT as a vaccine for preventing or treating pathological conditions from non-  
PT specific febrile illnesses to severe hemorrhagic manifestations or  
PT encephalitic syndromes.  
XX  
PS Disclosure; SEQ ID NO 35; 30pp; English.  
XX  
CC The present invention relates to an isolated and purified ApoptoM  
CC peptide. The invention is useful as a vaccine for the prevention and  
CC treatment of pathological conditions from non-specific febrile illnesses  
CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
CC pathological conditions are linked to Flavivirus infection or cancers.  
CC The invention is also useful in gene therapy. The present sequence is a  
CC p (95-114) EGFP (enhanced green fluorescent protein) (M1-M40)DEN (dengue)-2  
XX (136F) plasmid DNA encoded protein.

Query Match 98.2%; Score 54; DB 9; Length 48;

Best Local Similarity 88.9%; Pred. No. 0.031;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
:|||||  
Db 40 IETWFLRHP 48

RESULT 3  
ADW12595  
ID ADW12595 standard; peptide; 9 AA.

XX  
AC ADW12595;

XX  
DT 24-MAR-2005 (first entry)

XX  
DE M32-40/DEN-2 mutant protein #1.

XX  
KW Gene therapy; protein purification; virucide; cytostatic; vaccine;  
KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
KW DEN; dengue; mutant; mutein.



XX Dengue virus.  
 XX US2004266987-A1.  
 XX 30-DEC-2004.  
 XX 30-JUN-2003; 2003US-00608029.  
 XX 30-JUN-2003; 2003US-00608029.  
 XX (INSP ) INST PASTEUR.  
 XX Despres P, Catteau A;  
 XX WPI; 2005-047647/05.  
 XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
 PT as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.  
 XX Example 3; Fig 4; 30pp; English.  
 XX The present invention relates to an isolated and purified ApoptoM  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC M32-40/DEN (dengue)-2 mutant protein.  
 XX Sequence 9 AA;  
 XX Query Match 87.3%; Score 48; DB 9; Length 9;  
 XX Best Local Similarity 77.8%; Pred. No. 2.1e+06;  
 XX Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VETWFLRHP 9  
 DB :||| ||||  
 1 IETWILRHP 9  
 RESULT 4  
 ADW12594  
 ID ADW12594 standard; peptide; 21 AA.  
 XX AC ADW12594;  
 XX 24-MAR-2005 (first entry)  
 XX M20-40/DEN-2 mutant protein.  
 XX Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 XX hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 XX DEN; dengue; mutant; mutein.  
 XX Dengue virus.  
 XX US2004266987-A1.  
 XX 30-DEC-2004.  
 XX 30-JUN-2003; 2003US-00608029.  
 XX 30-JUN-2003; 2003US-00608029.  
 XX (INSP ) INST PASTEUR.  
 XX Despres P, Catteau A;  
 XX WPI; 2005-047647/05.  
 XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
 PT as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.  
 XX Example 3; Fig 4; 30pp; English.  
 XX The present invention relates to an isolated and purified ApoptoM  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC M32-40/DEN (dengue)-2 mutant protein.  
 XX Sequence 9 AA;  
 XX Query Match 87.3%; Score 48; DB 9; Length 9;  
 XX Best Local Similarity 77.8%; Pred. No. 2.1e+06;  
 XX Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VETWFLRHP 9  
 DB :||| ||||  
 1 IETWILRHP 9

PT New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
 PT as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.  
 XX Example 3; Fig 4; 30pp; English.  
 XX The present invention relates to an isolated and purified ApoptoM  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC M20-40/DEN (dengue)-2 mutant protein.  
 XX Sequence 21 AA;  
 XX Query Match 87.3%; Score 48; DB 9; Length 21;  
 XX Best Local Similarity 77.8%; Pred. No. 0.15;  
 XX Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VETWFLRHP 9  
 DB :||| ||||  
 13 IETWILRHP 21  
 RESULT 5  
 ADW12593  
 ID ADW12593 standard; peptide; 32 AA.  
 XX AC ADW12593;  
 XX 24-MAR-2005 (first entry)  
 XX M10-40/DEN-2 mutant protein.  
 XX Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 XX hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 XX DEN; dengue; mutant; mutein.  
 XX Dengue virus.  
 XX US2004266987-A1.  
 XX 30-DEC-2004.  
 XX 30-JUN-2003; 2003US-00608029.  
 XX 30-JUN-2003; 2003US-00608029.  
 XX (INSP ) INST PASTEUR.  
 XX Despres P, Catteau A;  
 XX WPI; 2005-047647/05.  
 XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
 PT as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.  
 XX Example 3; Fig 4; 30pp; English.  
 XX The present invention relates to an isolated and purified ApoptoM  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC M10-40/DEN (dengue)-2 mutant protein.  
 XX Sequence 32 AA;

```

Query Match      87.3%; Score 48; DB 9; Length 32;
Best Local Similarity 77.8%; Pred. No. 0.24;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 1 VETWFLRHP 9
DB 24 IETWILRHP 32

RESULT 6
ADM12576
ID ADM12576 standard; peptide; 39 AA.
XX AC ADM12576;
XX DT 24-MAR-2005 (first entry)
XX DE M1-40/DEN-2 protein.
XX KW Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX KW DEN; dengue.
XX OS Dengue virus.
XX PN US2004266987-A1.
XX PD 30-DEC-2004.
XX PF 30-JUN-2003; 2003US-00608029.
XX PR 30-JUN-2003; 2003US-00608029.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX DR WPI; 2005-047647/05.
XX PT New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
XX PT as a vaccine for preventing or treating pathological conditions from non-
XX PT specific febrile illnesses to severe hemorrhagic manifestations or
XX PT encephalitic syndromes.
XX PS Example 3; SEQ ID NO 23; 30pp; English.
XX CC The present invention relates to an isolated and purified ApoptoM
XX CC peptide. The invention is useful as a vaccine for the prevention and
XX CC treatment of pathological conditions from non-specific febrile illnesses
XX CC to severe hemorrhagic manifestations, encephalitic syndromes and these
XX CC pathological conditions are linked to Flavivirus infection or cancers.
XX CC The invention is also useful in gene therapy. The present sequence is a
XX CC M1-40/DEN (dengue)-2 protein.
XX SQ Sequence 39 AA;

Query Match      87.3%; Score 48; DB 9; Length 39;
Best Local Similarity 77.8%; Pred. No. 0.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
DB 31 IETWILRHP 39

RESULT 7
AAE17432
ID AAE17432 standard; peptide; 40 AA.
XX AC AAE17432;
XX DT 29-AUG-2003 (revised)
XX DT 18-APR-2002 (first entry)

```

```

XX DE Dengue (DEN)-2 virus M ectodomain.
XX KW Dengue virus; PRM glycoprotein; E glycoprotein; apoptosis; virucide;
XX KW cancer; flavivirus infection; cytostatic; DEN-2 M ectodomain.
XX OS Dengue virus; 2.
XX PN WO200196376-A2.
XX PD 20-DEC-2001.
XX PF 18-JUN-2001; 2001WO-IB001570.
XX PR 16-JUN-2000; 2000US-0212129P.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Courageot M, Deubel V, Catteau A;
XX DR WPI; 2002-139706/18.
XX PT Novel apoptosis inducing polypeptide fragments of Dengue virus-1 or 2 M
XX PT protein, useful for inducing apoptosis in a cell of a human patient
XX PS suffering from cancer or flavivirus infection.
XX PS Claim 9; Fig 12; 45pp; English.
XX CC The invention relates to pro-apoptotic fragments of the Dengue virus
XX CC (DEN) PRM and E glycoproteins, methods for screening molecules capable of
XX CC inducing apoptosis and methods of inducing apoptosis in a cell. The
XX CC invention particularly relates to DEN-1 M (a membrane protein anchored in
XX CC envelope surrounding the nucleocapsid of the virus) ectodomain sequences, of
XX CC Den-1-C amino acid sequence and DEN-2 M ectodomain sequence. Sequences of
XX CC the invention are useful for inducing apoptosis in a cell of a patient
XX CC suffering from cancer or flavivirus infection. They are also useful for
XX CC screening molecules which inhibit apoptosis. The present sequence is DEN-
XX CC 2 virus M ectodomain. (Updated on 29-AUG-2003 to standardise OS field)
XX SQ Sequence 40 AA;

Query Match      87.3%; Score 48; DB 5; Length 40;
Best Local Similarity 77.8%; Pred. No. 0.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
DB 32 IETWILRHP 40

RESULT 8
ADM12578
ID ADM12578 standard; peptide; 40 AA.
XX AC ADM12578;
XX DT 24-MAR-2005 (first entry)
XX DE M1-40/YF.17D (T34, I36, I37, H39) mutant protein.
XX KW Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX KW YF; yellow fever; mutant; mutein.
XX OS Yellow fever virus.
XX PN US2004266987-A1.
XX PD 30-DEC-2004.
XX PF 30-JUN-2003; 2003US-00608029.
XX PR 30-JUN-2003; 2003US-00608029.

```

```

XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX XX
XX DR WPI; 2005-047647/05.
XX PT
XX PT New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
XX PT as a vaccine for preventing or treating pathological conditions from non-
XX PT specific febrile illnesses to severe hemorrhagic manifestations or
XX PT encephalitic syndromes.
XX PS
XX PS Example 3; SEQ ID NO 25; 30pp; English.
XX CC
XX CC The present invention relates to an isolated and purified ApoptoM
XX CC peptide. The invention is useful as a vaccine for the prevention and
XX CC treatment of pathological conditions from non-specific febrile illnesses
XX CC to severe hemorrhagic manifestations, encephalitic syndromes and these
XX CC pathological conditions are linked to Flavivirus infection or cancers.
XX CC The invention is also useful in gene therapy. The present sequence is a
XX CC M1-40/YF (yellow fever).17D (T34, I36, I37, H39) mutant protein.
XX SQ
XX Query Match 87.3%; Score 48; DB 9; Length 40;
XX Best Local Similarity 77.8%; Pred. No. 0.3;
XX Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 VETWFLRHP 9
XX Db :|||||
XX 32 IETWILRHP 40
XX
XX RESULT 9
XX AAEL17433
XX ID AAEL17433 standard; protein; 48 AA.
XX AC
XX AC AAEL17433;
XX DT 18-APR-2002 (first entry)
XX XX
XX DE (95-114)EGFP(206-245)DEN-2 fusion protein.
XX KW
XX KW Dengue virus; prM glycoprotein; E glycoprotein; apoptosis; virucide;
XX KW cancer; flavivirus infection; cytostatic; EGFP; DEN-2 protein;
XX KW enhanced green fluorescent protein; fusion protein; M ectodomain.
XX XX
XX OS Dengue virus; 2.
XX OS Dengue virus; 1.
XX OS Unidentified.
XX OS Chimeric.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 13..44
XX FT /note= "Encoded by GTTATC"
XX XX
XX PN WO200196376-A2.
XX PN
XX PD 20-DEC-2001.
XX XX
XX PF 18-JUN-2001; 2001WO-IB001570.
XX XX
XX PR 16-JUN-2000; 2000US-0212129P.
XX XX
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Courageot M, Deubel V, Catteau A;
XX XX
XX DR WPI; 2002-139706/18.
XX DR N-PSDB; AAD27335.
XX XX
XX PT Novel apoptosis inducing polypeptide fragments of Dengue virus-1 or 2 M
XX PT protein, useful for inducing apoptosis in a cell of a human patient
XX
XX PT suffering from cancer or flavivirus infection.
XX XX
XX PS Claim 42; Fig 11; 45pp; English.
XX XX
XX CC The invention relates to pro-apoptotic fragments of the Dengue virus
XX CC (DEN) prM and E glycoproteins, methods for screening molecules capable of
XX CC inducing apoptosis and methods of inducing apoptosis in a cell. The
XX CC invention particularly relates to DEN-1 M (a membrane protein anchored in
XX CC envelope surrounding the nucleocapsid of the virus) ectodomain sequence.
XX CC Den-1-C amino acid sequence and DEN-2 M ectodomain sequence. Sequences of
XX CC the invention are useful for inducing apoptosis in a cell of a patient
XX CC suffering from cancer or flavivirus infection. They are also useful for
XX CC screening molecules which inhibit apoptosis. The present sequence is (95-
XX CC 114)EGFP(206-245)DEN-2 fusion protein construct. This construct comprises
XX CC 95-114 of the C-terminus of the C-protein of the DEN-1 virus strain BR/90
XX CC fused to the N-terminus of enhanced green fluorescent protein (EGFP) and
XX CC DEN-2 virus strain Jamaica M ectodomain (DEN-2 polypeptide) fused to the
XX CC C-terminus of the EGFP sequence
XX SQ
XX Sequence 48 AA;
XX
XX Query Match 87.3%; Score 48; DB 5; Length 48;
XX Best Local Similarity 77.8%; Pred. No. 0.37;
XX Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 VETWFLRHP 9
XX Db :|||||
XX 40 IETWILRHP 48
XX
XX RESULT 10
XX ADN37497
XX ID ADN37497 standard; protein; 167 AA.
XX AC
XX AC ADN37497;
XX DT 17-JUN-2004 (first entry)
XX XX
XX DE Dengue virus C15/truncated prM antigen fusion protein - SEQ ID 122.
XX KW
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;
XX KW capsid.
XX XX
XX OS Dengue virus.
XX XX
XX PN WO2003102166-A2.
XX XX
XX PD 11-DEC-2003.
XX XX
XX PF 26-FEB-2003; 2003WO-US005918.
XX XX
XX PR 26-FEB-2002; 2002US-0360030P.
XX XX
XX PA (MAXY-) MAXYGEN INC.
XX PI Apt D, Punnonen J, Brinkman AM;
XX XX
XX DR WPI; 2004-043106/04.
XX XX
XX PT New recombinant or synthetic polypeptides and polynucleotides useful for
XX PT diagnosing, preventing or treating diseases associated with flaviviruses,
XX PT including dengue viruses.
XX PS
XX PS Disclosure; SEQ ID NO 122; 409pp; English.
XX CC
XX CC The invention relates to a novel recombinant or synthetic polypeptide
XX CC comprising an amino acid sequence that has at least about 90% sequence
XX CC identity to any of the 20 fully defined amino acid sequences given in the
XX CC specification. The polypeptide of the invention demonstrates virucide
XX CC activity and may be useful for inducing an immune response to
XX CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
XX CC in detecting and/or diagnosing the presence of antibodies against the
XX CC Dengue virus serotypes in a sample and for gene therapy. The current

```

CC sequence is that of a Dengue virus C15/truncated prM antigen fusion  
 CC protein of the invention which comprises the C-terminal 15 amino acids of  
 CC the capsid protein fused to a truncated form of the prM protein lacking  
 CC the C-terminal 15 amino acids.

XX SQ Sequence 167 AA;

Query Match 87.3%; Score 48; DB 8; Length 167;  
 Best Local Similarity 77.8%; Pred. No. 1.4;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
 DB 139 IETWILRHP 147

RESULT 11

ADN37493  
 ID ADN37493 standard; protein; 171 AA.

XX AC ADN37493;

DT 17-JUN-2004 (first entry)

DE Dengue virus type 2 (DEN-2) C15/truncated prM antigen fusion protein.

XX virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;  
 KW capsid; DEN-2.

XX OS Dengue virus type 2.

PN WO2003102166-A2.

XX PD 11-DEC-2003.

PF 26-FEB-2003; 2003WO-US005918.

PR 26-FEB-2002; 2002US-0360030P.

PA (MAXY-) MAXYGEN INC.

PI Apt D, Punnonen J, Brinkman AM;

DR WPI; 2004-043106/04.

XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

PS Disclosure; SEQ ID NO 118; 409pp; English.

CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of a Dengue virus type 2 (DEN-2) C15/truncated prM  
 CC antigen fusion protein of the invention which comprises the C-terminal 15  
 CC amino acids of the capsid protein fused to a truncated form of the prM  
 CC protein lacking the C-terminal 15 amino acids.

XX SQ Sequence 171 AA;

Query Match 87.3%; Score 48; DB 8; Length 171;  
 Best Local Similarity 77.8%; Pred. No. 1.5;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
 DB 143 IETWILRHP 151

RESULT 12

ADN37496  
 ID ADN37496 standard; protein; 171 AA.

XX AC ADN37496;

DT 17-JUN-2004 (first entry)

DE Dengue virus C15/truncated prM antigen fusion protein - SEQ ID 121.

XX virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;  
 KW capsid.

XX OS Dengue virus.

PN WO2003102166-A2.

PD 11-DEC-2003.

PF 26-FEB-2003; 2003WO-US005918.

PR 26-FEB-2002; 2002US-0360030P.

PA (MAXY-) MAXYGEN INC.

PI Apt D, Punnonen J, Brinkman AM;

DR WPI; 2004-043106/04.

XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

PS Disclosure; SEQ ID NO 121; 409pp; English.

CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of a Dengue virus C15/truncated prM antigen fusion  
 CC protein of the invention which comprises the C-terminal 15 amino acids of  
 CC the capsid protein fused to a truncated form of the prM protein lacking  
 CC the C-terminal 15 amino acids.

XX SQ Sequence 171 AA;

Query Match 87.3%; Score 48; DB 8; Length 171;  
 Best Local Similarity 77.8%; Pred. No. 1.5;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
 DB 143 IETWILRHP 151

RESULT 13

AAW75410  
 ID AAW75410 standard; peptide; 635 AA.

XX AC AAW75410;

DT 17-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 02-MAR-1999 (first entry)

XX Fusion protein PD30 contains Dengue virus epitope.  
 DE  
 XX

KW Dengue virus; fusion protein; P64K; Neisseria meningitidis; epitope;  
 KW antibody; diagnosis; Flavivirus; infection; vaccine.  
 XX Dengue virus.  
 OS Neisseria meningitidis.  
 OS Chimeric.  
 PN WO9831814-A1.  
 XX  
 XX 23-JUL-1998.  
 XX  
 XX 13-JAN-1998; 98WO-CU000001.  
 XX  
 XX 15-JAN-1997; 97CU-00000013.  
 XX  
 XX (CIGB-) CIGB CENT ING GENETICA & BIOTECNOLOGIA.  
 PA (IPKM-) IPK INST MEDICINA TROPICAL KOURI PEDRO.  
 XX  
 XX Vazquez Ramado S, Guzman Tirado G, Guillen Nieto GE, Pardo Lazo OL;  
 PI Chinae Santiago G, Perez Diaz AB, Pupo Antunez M, Rodriguez Roche R;  
 PI Reyes Acosta O, Garay Perez HE, Padron Palomares G, Alvarez Vera M;  
 PI Morier Diaz L, Perez Insuaita O, Pelegrino Martinez De La Coterri Pedro;  
 DR WPI; 1998-414111/35.  
 XX  
 XX New peptide(s) and fusion proteins useful for diagnosis and treatment of  
 PT flavivirus infection - contain cross-reactive epitopes from Dengue virus  
 PT pre-M/M protein and can induce neutralising antibodies.  
 XX  
 XX Claim 7; Page 28-29; 64pp; Spanish.  
 XX  
 XX This protein represents a fusion protein comprising an M protein epitope  
 CC from Dengue virus type 2 inserted into the P64K protein from Neisseria  
 CC meningitidis. Synthetic peptides based on the Dengue virus epitope  
 CC sequences (AAW75404-W75408) and fusion proteins can be used to raise  
 CC antibodies. The peptides, protein and antibodies are all useful for  
 CC diagnosis and treatment of Flavivirus infection, e.g. in vaccines.  
 CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to  
 CC standardise OS field)  
 XX  
 XX Sequence 635 AA;  
 SQ  
 Query Match 87.3%; Score 48; DB 2; Length 635;  
 Best Local Similarity 77.8%; Pred. No. 6;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VETWFLRHP 9  
 Db :||| ||||  
 74 IETWILRHP 82  
 RESULT 14  
 ADN37628  
 ID ADN37628 standard; protein; 675 AA.  
 XX  
 XX AC ADN37628;  
 XX  
 XX 17-JUN-2004 (first entry)  
 XX  
 XX Dengue virus C15/prM/E part codon-optimised antigen fusion protein 2.  
 DE  
 XX  
 XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E;  
 KW human codon-optimised; prM; envelope; capsid.  
 XX  
 XX Dengue virus.  
 OS Synthetic.  
 OS  
 XX WO2003102166-A2.  
 PN  
 XX 11-DEC-2003.  
 PD  
 XX 26-FEB-2003; 2003WO-US005918.  
 XX  
 XX Claim 40; SEQ ID NO 143; 409pp; English.

PR 26-FEB-2002; 2002US-0360030P.  
 XX (MAXY-) MAXYGEN INC.  
 XX  
 XX Apt D, Punnonen J, Brinkman AM;  
 PI  
 XX WPI; 2004-043106/04.  
 DR N-PSDB; ADN37632.  
 DR  
 XX  
 XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX  
 XX Example 28; SEQ ID NO 253; 409pp; English.  
 PS  
 XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E partially human codon-  
 CC optimised antigen fusion protein of the invention which comprises 15  
 CC amino acids of the capsid (C) protein fused to the full-length partially  
 CC codon-optimised prM protein and envelope (E) protein.  
 XX  
 XX Sequence 675 AA;  
 SQ  
 Query Match 87.3%; Score 48; DB 8; Length 675;  
 Best Local Similarity 77.8%; Pred. No. 6.4;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VETWFLRHP 9  
 Db :||| ||||  
 139 IETWILRHP 147  
 RESULT 15  
 ADN37518  
 ID ADN37518 standard; protein; 675 AA.  
 XX  
 XX AC ADN37518;  
 XX  
 XX 17-JUN-2004 (first entry)  
 XX  
 XX Dengue virus C15/prM/E antigen fusion protein - SEQ ID 143.  
 DE  
 XX  
 XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;  
 KW envelope; capsid.  
 XX  
 XX Dengue virus.  
 OS  
 XX WO2003102166-A2.  
 PN  
 XX 11-DEC-2003.  
 PD  
 XX 26-FEB-2003; 2003WO-US005918.  
 PF  
 XX 26-FEB-2002; 2002US-0360030P.  
 PR  
 XX (MAXY-) MAXYGEN INC.  
 XX  
 XX Apt D, Punnonen J, Brinkman AM;  
 PI  
 XX WPI; 2004-043106/04.  
 DR  
 XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX  
 XX Claim 40; SEQ ID NO 143; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.  
 XX  
 XX Sequence 675 AA;  
 SQ

Query Match 87.3%; Score 48; DB 8; Length 675;  
 Best Local Similarity 77.8%; Pred. No. 6.4;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
 Db 139 IETWILRHP 147  
 :||| |||||

RESULT 16  
 ADN37612  
 ID ADN37612 standard; protein; 675 AA.  
 AC ADN37612;  
 DT 17-JUN-2004 (first entry)  
 DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 237.  
 XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;  
 KW envelope; capsid.  
 XX Dengue virus.  
 OS WO2003102166-A2.  
 PN 11-DEC-2003.  
 PD 26-FEB-2003; 2003WO-US005918.  
 PF 26-FEB-2002; 2002US-0360030P.  
 PR (MAXY-) MAXYGEN INC.  
 XX Apt D, Punnonen J, Brinkman AM;  
 PI WPI; 2004-043106/04.  
 DR New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX Claim 40; SEQ ID NO 237; 409pp; English.  
 PS The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.  
 XX  
 XX Sequence 675 AA;  
 SQ

Query Match 87.3%; Score 48; DB 8; Length 675;  
 Best Local Similarity 77.8%; Pred. No. 6.4;  
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 VETWFLRHP 9  
 Db 139 IETWILRHP 147  
 :||| |||||

RESULT 17  
 ADN37626  
 ID ADN37626 standard; protein; 675 AA.  
 AC ADN37626;  
 DT 17-JUN-2004 (first entry)  
 DE Dengue virus C15/prM/E part codon-optimised antigen fusion protein 1.  
 XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E;  
 KW human codon-optimised; prM; envelope; capsid.  
 XX Dengue virus.  
 OS Synthetic.  
 OS WO2003102166-A2.  
 PN 11-DEC-2003.  
 PD 26-FEB-2003; 2003WO-US005918.  
 PF 26-FEB-2002; 2002US-0360030P.  
 PR (MAXY-) MAXYGEN INC.  
 XX Apt D, Punnonen J, Brinkman AM;  
 PI WPI; 2004-043106/04.  
 DR N-PSDB; ADN37630.  
 XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX Claim 40; SEQ ID NO 251; 409pp; English.  
 PS The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E partially human codon-  
 CC optimised antigen fusion protein of the invention which comprises 15  
 CC amino acids of the capsid (C) protein fused to the full-length partially  
 CC codon-optimised prM protein and envelope (E) protein.  
 XX  
 XX Sequence 675 AA;  
 SQ

Query Match 87.3%; Score 48; DB 8; Length 675;  
 Best Local Similarity 77.8%; Pred. No. 6.4;  
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 VETWFLRHP 9  
 Db 139 IETWILRHP 147  
 :||| |||||

RESULT 18  
 AAW75411  
 ID AAW75411 standard; peptide; 677 AA.

```

XX AAW75411;
XX AC
XX DT 17-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 02-MAR-1999 (first entry)
XX DE
XX DE Fusion protein PD34 contains Dengue virus epitope.
XX KW
XX KW Dengue virus; fusion protein; P64K; Neisseria meningitidis; epitope;
XX KW antibody; diagnosis; Flavivirus; infection; vaccine.
XX OS
XX OS Dengue virus.
XX OS Neisseria meningitidis.
XX OS Chimeric.
XX PN
XX PN WO9831814-A1.
XX XX
XX PD 23-JUL-1998.
XX PF
XX PF 13-JAN-1998; 98WO-CU000001.
XX PR
XX PR 15-JAN-1997; 97CU-00000013.
XX PA
XX PA (CIGB-) CIGB CENT ING GENETICA & BIOTECNOLOGIA.
XX PA (IPKM-) IPK INST MEDICINA TROPICAL KOURI PEDRO.
XX PI
XX PI Vazquez Ramudo S, Guzman Tirado G, Guillen Nieto GE, Pardo Lazo OL;
XX PI Chinae Santiago G, Perez Diaz AB, Pupo Antunez M, Rodriguez Roche R;
XX PI Reyes Acosta O, Garay Perez HE, Padron Palomares G, Alvarez Vera M;
XX PI Morier Diaz L, Perez Insuaita O, Pelegrino Martinez De La Coterra Pedro;
XX XX
XX WPI; 1998-414111/35.
XX DR
XX DR New peptide(s) and fusion proteins useful for diagnosis and treatment of
XX PT flavivirus infection - contain cross-reactive epitopes from Dengue virus
XX PT pre-M/M protein and can induce neutralising antibodies.
XX PS
XX PS Claim 7; Page 30-32; 64pp; Spanish.
XX CC
XX CC This protein represents a fusion protein comprising an M protein epitope
XX CC from Dengue virus type 4 inserted into the P64K protein from Neisseria
XX CC meningitidis. Synthetic peptides based on the Dengue virus epitope
XX CC sequences (AAW75404-W75408) and fusion proteins can be used to raise
XX CC antibodies. The peptides, protein and antibodies are all useful for
XX CC diagnosis and treatment of flavivirus infection, e.g. in vaccines.
XX CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 677 AA;

Query Match 87.3%; Score 48; DB 2; Length 677;
Best Local Similarity 77.8%; Pred. No. 6.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db :||| ||||
116 IETWILRHP 124

RESULT 19
ADN37613
ID ADN37613 standard; protein; 677 AA.
XX AC
XX AC ADN37613;
XX DT
XX DT 17-JUN-2004 (first entry)
XX DE
XX DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 238.
XX KW
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;
XX KW envelope; capsid.

```

```

OS Dengue virus.
XX WO2003102166-A2.
XX PD
XX PD 11-DEC-2003.
XX PF
XX PF 26-FEB-2003; 2003WO-US005918.
XX PR
XX PR 26-FEB-2002; 2002US-0360030P..
XX PA
XX PA (MAXY-) MAXYGEN INC.
XX PI
XX PI Apt D, Punnonen J, Brinkman AM;
XX DR
XX DR WPI; 2004-043106/04.
XX PT
XX PT New recombinant or synthetic polypeptides and polynucleotides useful for
XX PT diagnosing, preventing or treating diseases associated with flaviviruses,
XX PT including dengue viruses.
XX PS
XX PS Example 13; SEQ ID NO 238; 409pp; English.
XX CC
XX CC The invention relates to a novel recombinant or synthetic polypeptide
XX CC comprising an amino acid sequence that has at least about 90% sequence
XX CC identity to any of the 20 fully defined amino acid sequences given in the
XX CC specification. The polypeptide of the invention demonstrates virucide
XX CC activity and may be useful for inducing an immune response to
XX CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
XX CC in detecting and/or diagnosing the presence of antibodies against the
XX CC Dengue virus serotypes in a sample and for gene therapy. The current
XX CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of
XX CC the invention which comprises 15 amino acids of the capsid (C) protein
XX CC fused to the full-length prM protein and envelope (E) protein.
XX SQ Sequence 677 AA;

Query Match 87.3%; Score 48; DB 8; Length 677;
Best Local Similarity 77.8%; Pred. No. 6.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db :||| ||||
139 IETWILRHP 147

RESULT 20
ADN37603
ID ADN37603 standard; protein; 681 AA.
XX AC
XX AC ADN37603;
XX DT
XX DT 17-JUN-2004 (first entry)
XX DE
XX DE Dengue virus type 2 Den-2C15/prM/E antigen fusion protein.
XX KW
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; DEN-2;
XX KW Den-2C15/prM/E; prM; envelope; capsid.
XX OS
XX OS Dengue virus type 2.
XX PN
XX PN WO2003102166-A2.
XX PD
XX PD 11-DEC-2003.
XX PF
XX PF 26-FEB-2003; 2003WO-US005918.
XX PR
XX PR 26-FEB-2002; 2002US-0360030P.
XX PA
XX PA (MAXY-) MAXYGEN INC.
XX PI
XX PI Apt D, Punnonen J, Brinkman AM;
XX DR
XX DR WPI; 2004-043106/04.

```

XX New recombinant or synthetic polypeptides and polynucleotides useful for  
PT diagnosing, preventing or treating diseases associated with flaviviruses,  
PT including dengue viruses.

XX Claim 38; SEQ ID NO 228; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
CC comprising an amino acid sequence that has at least about 90% sequence  
CC identity to any of the 20 fully defined amino acid sequences given in the  
CC specification. The polypeptide of the invention demonstrates virucide  
CC activity and may be useful for inducing an immune response to  
CC flaviviruses (arboviruses group B), including Dengue viruses, as well as  
CC in detecting and/or diagnosing the presence of antibodies against the  
CC Dengue virus serotypes in a sample and for gene therapy. The current  
CC sequence is that of the Dengue virus type 2 (DEN-2) Den-2C15/prM/E  
CC antigen fusion protein of the invention which comprises 15 amino acids of  
CC the capsid (C) protein fused to the full-length prM protein and envelope  
CC (E) protein.

XX Sequence 681 AA;

Query Match 87.3%; Score 48; DB 8; Length 681;  
Best Local Similarity 77.8%; Pred. No. 6.4;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
DB 143 IETWILRHP 151  
:|||||

RESULT 21  
ADN37517  
ID ADN37517 standard; protein; 681 AA.

XX ADN37517;  
AC  
XX  
XX 17-JUN-2004 (first entry)  
DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 142.

XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;  
KW envelope; capsid.  
XX Dengue virus.  
XX WO2003102166-A2.  
PN  
XX 11-DEC-2003.  
PD  
XX 26-FEB-2003; 2003WO-US005918.  
PF  
XX 26-FEB-2002; 2002US-0360030P.  
PR (MAXY-) MAXYGEN INC.  
XX  
XX Apt D, Punnonen J, Brinkman AM;  
PI  
XX WPI; 2004-043106/04.  
DR  
XX New recombinant or synthetic polypeptides and polynucleotides useful for  
PT diagnosing, preventing or treating diseases associated with flaviviruses,  
PT including dengue viruses.

XX Claim 40; SEQ ID NO 142; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
CC comprising an amino acid sequence that has at least about 90% sequence  
CC identity to any of the 20 fully defined amino acid sequences given in the  
CC specification. The polypeptide of the invention demonstrates virucide  
CC activity and may be useful for inducing an immune response to  
CC flaviviruses (arboviruses group B), including Dengue viruses, as well as  
CC in detecting and/or diagnosing the presence of antibodies against the

CC Dengue virus serotypes in a sample and for gene therapy. The current  
CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
CC the invention which comprises 15 amino acids of the capsid (C) protein  
CC fused to the full-length prM protein and envelope (E) protein.

XX Sequence 681 AA;

Query Match 87.3%; Score 48; DB 8; Length 681;  
Best Local Similarity 77.8%; Pred. No. 6.4;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
DB 143 IETWILRHP 151  
:|||||

RESULT 22  
ABP57874  
ID ABP57874 standard; protein; 685 AA.

XX ABP57874;  
AC  
XX 07-FEB-2003 (first entry)  
DT  
XX Plasmid pCBD2-14-6 containing dengue-2 virus prM and E.  
DE  
XX Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;  
KW pCBD2-14-6; dengue virus; DEN-2.  
KW  
XX Unidentified.  
OS Dengue-2 virus.  
OS Chimeric.  
OS  
XX WO200281754-A1.  
PN  
XX 17-OCT-2002.  
PD  
XX 04-APR-2002; 2002WO-US010764.  
PF  
XX 04-APR-2001; 2001US-00826115.  
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Chang GJ;  
PI  
XX WPI; 2003-058572/05.  
DR N-PSDB; ABV77547.  
DR  
XX Novel isolated nucleic acid useful as vaccine for preventing flavivirus  
PT infection, comprises transcriptional unit encoding signal sequence of one  
PT flavivirus and immunogenic flavivirus antigen of a second flavivirus.  
XX  
XX Example 20; Page 157-158; 174pp; English.

XX The invention relates to a novel nucleic acid comprising a  
CC transcriptional unit encoding a signal sequence of a structural protein  
CC of a first flavivirus and an immunogenic flavivirus antigen of a second  
CC flavivirus, where the transcriptional unit directs the synthesis of the  
CC antigen. The polynucleotide of the invention has virucide activity, and  
CC acts as a vaccine. A composition of the invention is useful for  
CC immunising a subject against infection by a flavivirus. The  
CC polynucleotide is useful as a vaccine for preventing flavivirus  
CC infection. The sequence represents plasmid pCBD2-14-6, which contains  
CC dengue-2 virus (DEN-2) prM and E proteins

XX Sequence 685 AA;

Query Match 87.3%; Score 48; DB 6; Length 685;  
Best Local Similarity 77.8%; Pred. No. 6.5;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
:|||||





```

FT Domain /label= Envelope
FT 296..395
FT /label= Domain-B
FT Misc-difference 588
FT /note= "amino acid residue 588 (Val) is Ile in wild-type
FT PR159"
FT Region 776..1127
FT /label= NS1
XX
XX WO9637221-A1.
XX
XX
XX 28-NOV-1996.
XX
XX 24-MAY-1996; 96WO-US007627.
XX
XX 24-MAY-1995; 95US-00448734.
XX 07-JUN-1995; 95US-00488807.
XX 10-JUL-1995; 95US-00500469.
XX
XX (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.
XX
XX Ivy JM, Nakano E, Clements D;
XX WPI; 1997-020938/02.
XX N-PSDB; AAT47666.
XX
XX Sub-unit vaccine against flavivirus infection - contg. recombinant
XX envelope protein in secretable form, used for immunising against
XX flavivirus infection.
XX
XX Example 1; Fig 3A-D; 121pp; English.
XX
XX A polypeptide (AAW09409) comprises the capsid, pre-membrane, envelope and
XX NS1 proteins of dengue virus serotype 2 (DEN-2) variant PR159/S1. A
XX conservative mutation in the envelope protein may be involved in the
XX attenuation of this small-plaque, temp.- sensitive variant. Portions of
XX the envelope protein, esp. domain B, can be expressed in eukaryotic hosts
XX (see also AAW09410 and AAW09427-28) transfected with vectors
XX incorporating DEN-2 S1 cDNA (see also AAT47666). These polypeptides can
XX be used in novel subunit vaccines against viral infection, to raise
XX antibodies useful for passive immunisation, and for diagnosis of
XX infection. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 1127 AA;
XX
XX Query Match 87.3%; Score 48; DB 2; Length 1127;
XX Best Local Similarity 77.8%; Pred. No. 11;
XX Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 VETWFLRHP 9
XX :||| ||||
XX 237 IETWILRHP 245
XX
XX RESULT 26
XX AAY05522
XX ID AAY05522 standard; protein; 1127 AA.
XX
XX AC AAY05522;
XX
XX 17-OCT-2003 (revised)
XX 05-JUL-1999 (first entry)
XX
XX Dengue virus serotype 2 PR159/S1 viral capsid, pprM, E, NS1.
XX
XX Flavivirus; envelope protein; vaccine; infection; diagnosis.
XX
XX Dengue virus; serotype 2.
XX
XX Key Location/Qualifiers
XX FT Protein 1..114
XX FT /label= Capsid
XX FT Protein 115..205

```

```

FT Protein /label= PreMembrane
FT 206..280
FT /label= Membrane
FT 280..1127
FT /label= Envelope
XX
XX WO9906068-A2.
XX
XX 11-FEB-1999.
XX
XX 27-JUL-1998; 98WO-US015447.
XX
XX 31-JUL-1997; 97US-00904227.
XX
XX (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.
XX
XX Ivy JM, Peters ID, Collier BG, McDonnell M, Harada KE;
XX WPI; 1999-153454/13.
XX N-PSDB; AAX25114.
XX
XX Recombinant dimeric flaviviral envelope vaccine - comprising a dimeric
XX 80% E protein, useful for protecting against flavivirus, especially dengue
XX virus infections.
XX
XX Example 1; Fig 3A-D; 60pp; English.
XX
XX This sequence is composed of the capsid, prM, envelope (E) and NS1
XX proteins of serotype 2 dengue virus DEN-2 strain PR159/S1. A vaccine for
XX protecting against flavivirus infection comprises a dimeric 80% E protein
XX that has been secreted as a recombinant protein from a eukaryotic cell.
XX 80% E indicates a C-terminally truncated flavivirus E protein. The
XX dimeric truncated E is formed: (1) by directly linking 2 tandem copies of
XX 80% E via a flexible tether; (2) via the formation of a leucine zipper
XX domain through the homodimeric association of 2 leucine zipper helices
XX each fused to the C-terminus of an 80% E molecule; or (3) via the
XX formation of a non-covalently associated four-helix bundle domain formed
XX upon association of two helix-turn-helix moieties attached to the C-
XX terminus of an 80% E molecule. Dimeric truncated DEN-2 E proteins are
XX efficiently secreted by recombinant cells, are easier to purify than
XX intracellular proteins, and generate a high titer neutralising antibody
XX response. The method is generally applicable to flaviviruses, in
XX particular dengue viruses such as DEN-2, where 80% E comprises amino
XX acids 1-395 of DEN-2 E. The products can also be used for diagnosis of
XX infection. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 1127 AA;
XX
XX Query Match 87.3%; Score 48; DB 2; Length 1127;
XX Best Local Similarity 77.8%; Pred. No. 11;
XX Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 VETWFLRHP 9
XX :||| ||||
XX 237 IETWILRHP 245
XX
XX RESULT 27
XX ADL98086
XX ID ADL98086 standard; protein; 1127 AA.
XX
XX AC ADL98086;
XX
XX 18-NOV-2004 (first entry)
XX
XX Dengue virus, DEN-2, capsid/membrane/envelope/NS1 proteins.
XX
XX Dengue virus; DEN-2; Envelope protein; 80% E; membrane protein;
XX capsid protein; NS1 protein; Dengue haemorrhagic fever; DHF;
XX Dengue shock syndrome; DSS; flavivirus; vaccine.
XX
XX Dengue virus type 2; strain PR159/S1.
XX

```

PN US2003175304-A1.  
 XX 18-SEP-2003.  
 XX 20-SEP-2002; 2002US-00247960.  
 XX 31-JUL-1997; 97US-00904227.  
 PR 18-AUG-1999; 99US-00376463.  
 XX (PETE/) PETERS I D.  
 PA (COLL/) COLLIER B G.  
 PA (MCDO/) MCDONELL M.  
 PA (IVY/) IVY J M.  
 PA (HARA/) HARADA K.  
 XX Peters ID, Collier BG, McDonnell M, Ivy JM, Harada K;  
 PI WPI; 2003-898503/82.  
 XX N-PSDB; ADL98085.  
 DR Vaccine useful for protection against dengue virus infection, comprises a  
 PT dimeric 80% envelope, which has been secreted as a recombinantly produced  
 PT protein from Drosophila Schneider cells.  
 XX Example 1; Fig 3; 3lpp; English.  
 XX The invention relates to a vaccine for protection against Flavivirus  
 CC infection comprising a dimeric 80% envelope (E), which has been secreted  
 CC as a recombinantly produced protein from Drosophila Schneider cells and  
 CC which represents the N-terminal 80% portion of the protein from residue 1  
 CC -395. Also included are a method for protecting a subject against a  
 CC Flavivirus, an immunogenic polypeptide comprising a dimeric 80% E, an  
 CC immunogenic composition for protection against Flavivirus infection  
 CC comprising the immunogenic polypeptide and a carrier, an immunodiagnostic  
 CC for detecting Flavivirus comprising the immunogenic polypeptide, a vector  
 CC host recombinant DNA expression system, a DNA sequence encoding the  
 CC immunogenic polypeptide and an immunodiagnostic kit for detecting  
 CC Flavivirus in a test subject. The dimeric 80% E products are envelope  
 CC proteins of serotypes comprising DEN-1, DEN-2, DEN-3 or DEN-4. The  
 CC Flavivirus is a dengue virus. The 80% E protein is produced as a dimer by  
 CC incorporating 2 different kinds of leucine zipper peptides or  
 CC incorporating a helix-turn-helix peptide, to encourage dimerisation. The  
 CC vaccine is useful for protection against dengue virus infection (e.g.  
 CC Dengue haemorrhagic fever, DHF, and Dengue shock syndrome, DSS). The  
 CC present sequence is encoded by the partial genomic sequence of the DEN-2  
 CC strain PR159/S1 virus, and represents the capsid, membrane, envelope and  
 CC NS1 proteins.  
 XX Sequence 1127 AA;  
 SQ Query Match 87.3%; Score 48; DB 7; Length 1127;  
 Best Local Similarity 77.8%; Pred. No. 11;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VETWFLRHP 9  
 Db :||| ||||  
 237 IETWILRHP 245  
 RESULT 28  
 ADQ28716  
 ID ADQ28716 standard; protein; 1127 AA.  
 XX AC ADQ28716;  
 XX 26-AUG-2004 (first entry)  
 DE Dengue virus viral capsid, prM, E and NS1 gene polyprotein.  
 DE virucide; vaccine; Flavivirus; dimeric 80%E; Drosophila Schneider cell;  
 KW immunogenic composition; multivalent immunodiagnostic; dengue virus;  
 KW viral capsid; prM gene; E gene; NS1 gene.  
 XX

OS Dengue virus.  
 XX US6749857-B1.  
 XX 15-JUN-2004.  
 XX 18-AUG-1999; 99US-00376463.  
 PF 31-JUL-1997; 97US-00904227.  
 PR (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.  
 PA Peters ID, Collier BG, McDonnell M, Ivy JM, Harada K;  
 PI WPI; 2004-438725/41.  
 XX N-PSDB; ADQ28715.  
 DR New vaccines for preventing or diagnosing infections caused by dengue  
 PT virus comprises a therapeutic amount of a dimeric 80%E protein secreted  
 PT from Drosophila Schneider cells.  
 XX Example 1; SEQ ID NO 3; 47pp; English.  
 XX The invention describes a vaccine that generates a protective,  
 CC neutralising antibody response to a Flavivirus in a murine host. The  
 CC vaccine comprises a therapeutic amount of a dimeric 80%E, the dimeric  
 CC 80%E having been secreted as a recombinantly produced protein from  
 CC Drosophila Schneider cells, and where 80%E represents the N-terminal 80%  
 CC portion of the protein from residues 1-395. Also described are: an  
 CC immunogenic polypeptide comprising the dimeric 80%E cited above; an  
 CC immunogenic composition that generates a protective, neutralising  
 CC antibody response to a Flavivirus in a murine host, comprising the above  
 CC immunogenic polypeptide and a physiological carrier; a multivalent  
 CC immunodiagnostic for the detection of Flavivirus, comprising at least 2  
 CC of the above immunogenic polypeptides of at least 2 flaviviral serotypes;  
 CC and an immunodiagnostic kit for the detection of Flavivirus in a test  
 CC subject, comprising the above immunogenic or multivalent immunodiagnostic  
 CC polypeptide, a suitable support phase coated with dimeric 80%E, and  
 CC labeled antibodies immunoreactive to antibodies from the test subject.  
 CC The composition is useful for preventing or diagnosing infections caused  
 CC by dengue virus. This is the amino acid sequence of the polyprotein  
 CC encoded by Dengue virus gene viral capsid, prM, E and NS1 genes for  
 CC Dengue virus strain PR159/S1 used as the source of DEN-2 genes for the  
 CC invention.  
 XX Sequence 1127 AA;  
 SQ Query Match 87.3%; Score 48; DB 8; Length 1127;  
 Best Local Similarity 77.8%; Pred. No. 11;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VETWFLRHP 9  
 Db :||| ||||  
 237 IETWILRHP 245  
 RESULT 29  
 AAE35314  
 ID AAE35314 standard; protein; 3388 AA.  
 XX AC AAE35314;  
 XX 28-MAY-2003 (first entry)  
 DE Dengue virus type 2 strain rDEN2/delta30 protein.  
 DE Attenuation; growth; vaccine; infection; Dengue virus type 4.  
 KW Dengue virus.  
 OS WO200295075-A1.  
 XX 28-NOV-2002.  
 PD

XX 22-MAY-2002; 2002WO-US016308.  
 PF 22-MAY-2001; 2001US-0293049P.  
 PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (BLAN/) BLANEY J E.  
 XX whitehead SS, Murphy BR, Hanley KA;  
 PI WPI; 2003-120809/11.  
 DR N-PSDB; AAD53912.  
 XX New mutated flavivirus, useful for fine tuning the attenuation and growth  
 PT characteristics of dengue virus vaccines for the prevention and/or  
 PT treatment of dengue virus infection.  
 XX Disclosure; Page 133-134; 246pp; English.  
 XX The present invention relates to novel mutated flaviviruses comprising a  
 CC phenotype in which the viral genome is modified by introduction of a  
 CC mutation, singly or in combination, taken from mutations from recombinant  
 CC virus bearing Vero adaptation mutations, putative Vero cell adaptation  
 CC mutations of dengue type 4 virus (DEN4) or mutations known to attenuate  
 CC dengue type 4 virus. The methods and compositions of the invention are  
 CC useful for fine tuning the attenuation and growth characteristics of  
 CC dengue virus vaccines for the prevention and/or treatment of dengue virus  
 CC infection. The present sequence is Dengue virus type 4 strain  
 CC rDEN2/4delta30 protein  
 XX Sequence 3388 AA;  
 SQ

Query Match 87.3%; Score 48; DB 6; Length 3388;  
 Best Local Similarity 77.8%; Pred. No. 36;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VETWFLRHP 9  
 DB 237 IETWILRHP 245

RESULT 30  
 AAR13166  
 ID AAR13166 standard; protein; 3391 AA.  
 XX AAR13166;  
 AC 25-MAR-2003 (revised)  
 DT 21-NOV-1991 (first entry)  
 XX Proteins encoded by entire Dengue 2 virus genome.  
 DE dengue virus; detection; consensus sequence; Flavivirus; PCR.  
 XX Dengue virus.  
 OS Key Location/Qualifiers  
 XX Peptide 116..205  
 FT /label= prM  
 FT Modified-site 183  
 FT /label= N-glycosylated  
 FT Protein 206..280  
 FT /label= M  
 FT Protein 281..775  
 FT /label= E  
 FT Modified-site 347  
 FT /label= N-glycosylated  
 FT Modified-site 433  
 FT /label= N-glycosylated  
 FT Protein 776..1127  
 FT Modified-site 905  
 FT /label= NS1  
 FT Modified-site 905  
 FT /label= N-glycosylated

FT Modified-site 982  
 FT /label= N-glycosylated  
 FT Protein 1128..1345  
 FT /label= NS2A  
 FT Modified-site 1134  
 FT /label= N-glycosylated  
 FT Modified-site 1174  
 FT /label= N-glycosylated  
 FT Modified-site 1329  
 FT /label= N-glycosylated  
 FT Protein 1346..1474  
 FT /label= NS2B  
 FT Modified-site 1369  
 FT /label= N-glycosylated  
 FT Protein 1475..2093  
 FT /label= NS3  
 FT Protein 2094..2243  
 FT /label= ns4a  
 FT Protein 2244..2492  
 FT /label= NS4B  
 FT Modified-site 2301  
 FT /label= N-glycosylated  
 FT Modified-site 2305  
 FT /label= N-glycosylated  
 FT Modified-site 2457  
 FT /label= N-glycosylated  
 FT Modified-site 2485  
 FT /label= N-glycosylated  
 FT Protein 2493..3391  
 FT /label= NS5  
 FT Modified-site 2644  
 FT /label= N-glycosylated  
 FT Modified-site 2665  
 FT /label= N-glycosylated  
 FT Modified-site 2704  
 FT /label= N-glycosylated  
 FT Modified-site 2714  
 FT /label= N-glycosylated  
 XX FR2654113-A.  
 PN 10-MAY-1991.  
 PD 09-NOV-1989; 89FR-00914724.  
 PF 09-NOV-1989; 89FR-00014724.  
 XX (INSP ) INST PASTEUR.  
 XX Vincent D;  
 PI WPI; 1991-225002/31.  
 DR N-PSDB; AAQ12787.  
 XX Detection and identification of Flaviviridae in biological sample - by  
 PT amplifying consensus sequence then hybridisation opt. followed by typing,  
 PT e.g. sequencing amplified prod.  
 XX Disclosure; Fig 3; 24pp; French.  
 CC The dengue 2 virus is an example of a member of the Flaviviridae which  
 CC can be identified using the probe pair of the invention. A species-  
 CC specific sequence can be amplified using the claimed oligonucleotides as  
 CC primers in a PCR reaction (see AAQ12788 and AAQ12789). Other viruses  
 CC which can be identified include Japanese encephalitis virus and yellow  
 CC fever virus. All the dengue 2 virus proteins are encoded from an  
 CC uninterrupted genomic sequence. (Updated on 25-MAR-2003 to correct PR  
 CC field.)  
 XX Sequence 3391 AA;  
 SQ

Query Match 87.3%; Score 48; DB 2; Length 3391;  
 Best Local Similarity 77.8%; Pred. No. 36;

---

Matches	7;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	VETWFLRHP	9						
		:							
Db	237	IETWILRHP	245						

Search completed: August 31, 2006, 11:50:37  
Job time : 111.25 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 11:43:31 ; Search time 17.25 Seconds  
(without alignments)  
50.200 Million cell updates/sec

Title: DENGUE\_SEROTYPE1  
Perfect score: 55  
Sequence: 1 vetflrhp 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 80:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	87.3	166	2 S40144	premembrane protein
2	48	87.3	555	2 JQ1404	genome polyprotein
3	48	87.3	775	2 A48644	polyprotein - deng
4	48	87.3	3388	1 GNVWDP	genome polyprotein
5	48	87.3	3391	1 GNVW16	genome polyprotein
6	48	87.3	3391	1 GNVW26	genome polyprotein
7	48	87.3	3391	1 GNVWJA	genome polyprotein
8	48	87.3	3391	2 JS0219	polyprotein - deng
9	47	85.5	555	2 JQ1405	genome polyprotein
10	47	85.5	775	2 A47311	polyprotein(C, E,
11	47	85.5	792	2 C32401	genome polyprotein
12	47	85.5	792	2 B32401	genome polyprotein
13	47	85.5	792	2 A32401	genome polyprotein
14	47	85.5	1226	1 GNVWTP	genome polyprotein
15	47	85.5	3390	1 GNVWD3	genome polyprotein
16	47	85.5	3396	1 A42551	genome polyprotein
17	46	83.6	1127	1 GNVWD2	genome polyprotein
18	45	81.8	166	2 S09223	membrane protein -
19	45	81.8	166	2 S09225	membrane protein -
20	42	76.4	665	2 PS0043	genome polyprotein
21	41	74.5	166	2 S09224	membrane protein -
22	41	74.5	422	2 A83184	probable protein m
23	40	72.7	205	2 E86085	hypothetical prote
24	40	72.7	205	2 A98238	hypothetical prote
25	39	70.9	205	2 I78665	hypothetical 23.0K
26	39	70.9	343	2 H95879	probable sugar ABC
27	38	69.1	144	2 B40098	colorectal cancer
28	38	69.1	773	2 A47666	structural polypro
29	38	69.1	1155	2 B96761	probable protein k

30	38	69.1	1244	2 S37034	DNA-directed DNA p
31	38	69.1	1447	2 A54100	tumor suppressor p
32	38	69.1	1525	1 GNVV55	genome polyprotein
33	38	69.1	3386	1 GNVWDF	genome polyprotein
34	38	69.1	3411	1 GNVVY	genome polyprotein
35	38	69.1	3411	1 GNVVYP	genome polyprotein
36	37	67.3	266	2 S02510	nifM protein - Kle
37	37	67.3	301	2 C95872	hypothetical prote
38	37	67.3	399	2 T49934	carboxypeptidase-1
39	37	67.3	427	2 F72389	conserved hypotet
40	37	67.3	533	2 T35722	probable transport
41	37	67.3	560	1 VGBE14	glycoprotein gpv -
42	37	67.3	640	2 B32935	hypothetical prote
43	37	67.3	826	2 B96712	probable receptor
44	37	67.3	2413	2 S34670	splicing factor PR
45	36	65.5	217	2 A83146	lipote-protein li
46	36	65.5	267	2 A38442	probable tumor sup
47	36	65.5	343	2 G84711	hypothetical prote
48	36	65.5	417	1 VGBE1B	glycoprotein D pre
49	36	65.5	436	2 S35784	glycoprotein gp -
50	36	65.5	481	2 H69588	acetylornithine de
51	36	65.5	490	2 E83062	deoxyribodipyrimid
52	36	65.5	575	2 I41293	EcoE type I restri
53	36	65.5	575	2 A49667	interleukin-10 rec
54	36	65.5	879	2 B70014	hypothetical prote
55	36	65.5	1008	2 T12532	antibiotic synthet
56	36	65.5	1008	2 T12532	hypothetical prote
57	36	65.5	2236	1 QZ5F	rudimentary protei
58	35	63.6	120	2 A97655	hypothetical prote
59	35	63.6	120	2 AG2878	conserved hypotet
60	35	63.6	194	1 S49184	phosphinothricin N
61	35	63.6	295	2 AG0923	LysR-family regula
62	35	63.6	297	2 F98323	hypothetical oxido
63	35	63.6	297	2 AH2959	hypothetical dehyd
64	35	63.6	305	2 G84140	aryl-alcohol dehyd
65	35	63.6	306	2 B97315	aldol/keto reductas
66	35	63.6	328	2 E83321	conserved hypotet
67	35	63.6	336	2 JE0215	nitrite reductase
68	35	63.6	360	2 JG0170	nitrite reductase
69	35	63.6	384	2 S74774	hypothetical prote
70	35	63.6	489	2 A47200	EcoA system protei
71	35	63.6	493	2 F86133	hypothetical prote
72	35	63.6	493	2 C91292	hypothetical prote
73	35	63.6	516	1 FWSYG3	glycinin G5 precur
74	35	63.6	630	2 T02524	probable RING zinc
75	35	63.6	739	2 A90141	ATP-dependent heli
76	35	63.6	805	2 G87268	DNA gyrase subunit
77	35	63.6	815	2 T41490	hypothetical prote
78	35	63.6	1310	2 T40135	oxysterol-binding
79	35	63.6	1332	2 F69732	PBSX prophage ORF
80	35	63.6	1467	2 T23950	hypothetical prote
81	35	63.6	1693	2 AC3240	helicase, SNF2 fam
82	35	63.6	4196	2 T43274	dyslin heavy chain
83	34	61.8	185	2 D83435	conserved hypotet
84	34	61.8	208	2 T33341	hypothetical prote
85	34	61.8	216	2 H72291	hypothetical prote
86	34	61.8	224	2 B87657	conserved hypotet
87	34	61.8	225	2 G72291	hypothetical prote
88	34	61.8	235	2 AF0656	conserved hypotet
89	34	61.8	244	2 D84979	phosphoadenosine p
90	34	61.8	246	2 H70223	conserved hypotet
91	34	61.8	256	2 F83223	conserved hypotet
92	34	61.8	270	2 E64924	hypothetical prote
93	34	61.8	270	2 D85774	hypothetical prote
94	34	61.8	270	2 H90925	hypothetical prote
95	34	61.8	314	2 C81735	tRNA Delta-2-isope
96	34	61.8	322	2 E84908	hypothetical prote
97	34	61.8	333	2 T02690	hypothetical prote
98	34	61.8	336	2 C82146	probable tetraacyl
99	34	61.8	366	2 B36919	hypothetical prote
100	34	61.8	403	2 S42532	hypothetical prote

## ALIGNMENTS

```
RESULT 1
S40144
premembrane protein - dengue virus type 2
C:Species: dengue virus type 2
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 31-Dec-2004
C:Accession: S40144
R:Shiu, S.Y.W.
submitted to the EMBL Data Library, May 1993
A:Reference number: S40144
A:Accession: S40144
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-166 <SHI>
A:Cross-references: UNIPROT:Q66346; UNIPARC:UPI00000F6DD9; EMBL:X72849; NID:g437772; PID:
C:Superfamily: hepatitis C virus genome polyprotein

Query Match      87.3%; Score 48; DB 2; Length 166;
Best Local Similarity 77.8%; Pred. No. 0.36;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
DB 123 IETWILRHP 131

RESULT 2
JQ1404
genome polyprotein - dengue virus type 2 (strain TH-36) (fragment)
N:Contains: envelope protein E; membrane-associated protein M; nonstructural protein NS1
C:Species: dengue virus type 2
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004
C:Accession: JQ1404
R:Shiu, S.Y.W.; Jiang, W.R.; Porterfield, J.S.; Gould, E.A.
J. Gen. Virol. 73, 207-212, 1992
A:Title: Envelope protein sequences of dengue virus isolates TH-36 and TH-Sman, and ident
A:Reference number: JQ1404; MUID:92113574; PMID:1339466
A:Accession: JQ1404
A:Molecule type: genomic RNA
A:Residues: 1-555 <SHI>
A:Cross-references: UNIPROT:P29984; UNIPARC:UPI000131DF8; GB:D10514; DBBJ:D01074; NID:9
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: envelope protein; glycoprotein; nonstructural protein; polyprotein; transmem
F:1-45/Product: membrane-associated protein M (fragment) #status predicted <MEM>
F:37-53/Domain: transmembrane #status predicted <TM1>
F:50-544/Product: envelope protein E #status predicted <ENV>
F:496-512/Domain: transmembrane #status predicted <TM2>
F:526-542/Domain: transmembrane #status predicted <TM3>
F:545-555/Product: nonstructural protein NS1 (fragment) #status predicted <NON>
F:116,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      87.3%; Score 48; DB 2; Length 555;
Best Local Similarity 77.8%; Pred. No. 1.2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
DB 6 IETWILRHP 14

RESULT 3
A48644
polyprotein - dengue virus type 2 (strain Mexican) (fragment)
C:Species: dengue virus type 2
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 31-Dec-2004
C:Accession: A48644
R:Ruiz, B.H.; Sanchez, I.; Ortega, G.J.; Lopez, I.; Ortiz-Ortiz, L.
submitted to Genbank, October 1992
A:Description: Nucleotide sequence and deduced amino-acid sequence of the structural pro
A:Reference number: A48644
A:Accession: A48644
```

```
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-775 <RUI>
A:Cross-references: UNIPROT:Q66398; UNIPARC:UPI00000EEB45; GB:L04561; NID:g323652; PIDN:P
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein

Query Match      87.3%; Score 48; DB 2; Length 775;
Best Local Similarity 77.8%; Pred. No. 1.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
DB 237 IETWILRHP 245

RESULT 4
GNWVDP
genome polyprotein - dengue virus type 2 (strain PR159/S1)
N:Contains: capsid protein; envelope protein; membrane protein; nonstructural protein NS1
A: nonstructural protein NS4b; nonstructural protein NS5
C:Species: dengue virus type 2
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 31-Dec-2004
C:Accession: A29972
R:Hahn, Y.S.; Galler, R.; Hunkapiller, T.; Dalrymple, J.M.; Strauss, J.H.; Strauss, E.G.
Virology 162, 167-180, 1988
A:Title: Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with
A:Reference number: A29972; MUID:88101365; PMID:2827375
A:Accession: A29972
A:Molecule type: genomic RNA
A:Residues: 1-3188 <HAH>
A:Cross-references: UNIPARC:UPI0000131DFB; GB:M19197; NID:g323654; PIDN:AAA42962.1; PID:
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; membrane protein; nonst
F:2-114/Product: capsid protein #status predicted <CAP>
F:115-280/Product: membrane protein precursor #status predicted <MPP>
F:115-205/Domain: nonterminal signal sequence #status predicted <SIG>
F:206-280/Product: membrane protein #status predicted <MMP>
F:281-775/Product: envelope protein #status predicted <ENP>
F:776-1188/Product: nonstructural protein NS1 #status predicted <NS1>
F:1189-1345/Product: nonstructural protein NS2a #status predicted <N2A>
F:1346-1475/Product: nonstructural protein NS2b #status predicted <N2B>
F:1476-2090/Product: nonstructural protein NS3 #status predicted <NS3>
F:1668-1675/Region: nucleotide-binding motif A (P-loop)
F:1755-1760/Region: nucleotide-binding motif B
F:1759-1762/Region: DBAH motif
F:2091-2376/Product: nonstructural protein NS4a #status predicted <N4A>
F:2377-2488/Product: nonstructural protein NS4b #status predicted <N4B>
F:2489-3388/Product: nonstructural protein NS5 #status predicted <NS5>
F:183,347,433,905,982,1134,1174,1329,1369,2298,2302,2384,2454,2482,2641,2662,2701,2711/B;

Query Match      87.3%; Score 48; DB 1; Length 3388;
Best Local Similarity 77.8%; Pred. No. 7.6;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
DB 237 IETWILRHP 245

RESULT 5
GNWV16
genome polyprotein - dengue virus type 2 (strain 16681)
N:Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstruc
tural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: dengue virus type 2
C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 31-Dec-2004
C:Accession: A42451; A43496; A43763
R:Blak, J.; McWilliam, S.M.; Butler, H.C.; Gibbs, A.J.; Weiller, G.; Herring, B.L.; Hems;
Virology 187, 573-590, 1992
A:Title: Comparison of a dengue-2 virus and its candidate vaccine derivative: sequence re
A:Reference number: A42451; MUID:92189532; PMID:1312269
A:Accession: A42451
```



A;Molecule type: genomic RNA  
A;Residues: 1-3391 <BLO>  
A;Cross-references: UNIPROT:P29990; UNIPARC:UPI0000131DF5; GB:M85259; NID:g32  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F;1-114/Product: capsid protein C #status predicted <CPC>  
F;115-280/Product: membrane-associated protein M precursor #status predicted <MPP>  
F;115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
F;206-280/Product: membrane-associated protein M #status predicted <MPM>  
F;268-284/Domain: transmembrane #status predicted <TM1>  
F;281-775/Product: envelope protein E #status predicted <EPE>  
F;727-743/Domain: transmembrane #status predicted <TM2>  
F;757-773/Domain: transmembrane #status predicted <TM3>  
F;776-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
F;1128-1345/Product: nonstructural protein NS2a #status predicted <N2A>  
F;1346-1474/Product: nonstructural protein NS2b #status predicted <N2B>  
F;1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
F;1668-1675/Region: nucleotide-binding motif A (P-loop)  
F;1755-1760/Region: nucleotide-binding motif B  
F;1759-1762/Region: DEAH motif  
F;2094-2243/Product: nonstructural protein NS4a #status predicted <N4A>  
F;2244-2491/Product: nonstructural protein NS5 #status predicted <NS5>  
F;2492-3391/Product: nonstructural protein NS5 (covalent) #status predicted  
F;193,347,433/Binding site: carbohydrate (Asn) #status predicted

Query Match 87.3%; Score 48; DB 1; Length 3391;  
Best Local Similarity 77.8%; Pred. No. 7.6;  
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 VETWFLRHP 9  
:|||||  
Db 237 IETWILRHP 245

RESULT 6  
GNMW26  
genome polyprotein - dengue virus type 2 (strain 16681-PDK53)  
N;Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru  
tural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: dengue virus type 2  
C;Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 31-Dec-2004  
C;Accession: B42451  
R;Blok, J.; McWilliams, S.M.; Butler, H.C.; Gibbs, A.J.; Weiller, G.; Herring, B.L.; Heng  
Virology 187, 573-590, 1992  
A;Title: Comparison of a dengue-2 virus and its candidate vaccine derivative: sequence  
A;Reference number: A42451; MUID:92188532; PMID:1312269  
A;Accession: B42451  
A;Molecule type: genomic RNA  
A;Residues: 1-3391 <BLO>  
A;Cross-references: UNIPROT:P29991; UNIPARC:UPI0000131DF6; GB:M85259  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F;1-114/Product: capsid protein C #status predicted <CPC>  
F;50-66/Domain: transmembrane #status predicted <TM1>  
F;102-118/Domain: transmembrane #status predicted <TM2>  
F;115-280/Product: membrane-associated protein M precursor #status predicted <MPP>  
F;115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
F;206-280/Product: membrane-associated protein M #status predicted <MPM>  
F;268-284/Domain: transmembrane #status predicted <TM3>  
F;281-775/Product: envelope protein E #status predicted <EPE>  
F;727-743/Domain: transmembrane #status predicted <TM4>  
F;757-773/Domain: transmembrane #status predicted <TM5>  
F;776-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
F;1128-1345/Product: nonstructural protein NS2a #status predicted <N2A>  
F;1158-1174/Domain: transmembrane #status predicted <TM6>  
F;1272-1288/Domain: transmembrane #status predicted <TM7>  
F;1294-1310/Domain: transmembrane #status predicted <TM8>  
F;1346-1474/Product: nonstructural protein NS2b #status predicted <N2B>  
F;1351-1367/Domain: transmembrane #status predicted <TM9>  
F;1373-1389/Domain: transmembrane #status predicted <TMA>  
F;1448-1464/Domain: transmembrane #status predicted <TM8>  
F;1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
F;1668-1675/Region: nucleotide-binding motif A (P-loop)

F;1755-1760/Region: nucleotide-binding motif B  
F;1759-1762/Region: DEAH motif  
F;2094-2243/Product: nonstructural protein NS4a #status predicted <N4A>  
F;2148-2164/Domain: transmembrane #status predicted <TM2>  
F;2174-2190/Domain: transmembrane #status predicted <TM3>  
F;2197-2213/Domain: transmembrane #status predicted <TM3>  
F;2227-2243/Domain: transmembrane #status predicted <TMF>  
F;2244-2491/Product: nonstructural protein NS4b #status predicted <N4B>  
F;2352-2368/Domain: transmembrane #status predicted <TMG>  
F;2411-2427/Domain: transmembrane #status predicted <TMH>  
F;2492-3391/Product: nonstructural protein NS5 #status predicted <NS5>  
F;183,347,433,905,982,1134,1174,1329,2301,2305,2346,2387,2457,2485,2644,2665,2704,2714/B  
Query Match 87.3%; Score 48; DB 1; Length 3391;  
Best Local Similarity 77.8%; Pred. No. 7.6;  
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 VETWFLRHP 9  
:|||||  
Db 237 IETWILRHP 245

RESULT 7  
GNMWJA  
genome polyprotein - dengue virus type 2 (strain Jamaica)  
N;Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru  
tural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: dengue virus type 2  
C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 31-Dec-2004  
C;Accession: A94346; A94378; A25613; A29199  
R;Deubel, V.; Kinney, R.M.; Trent, D.W.  
Virology 155, 365-377, 1986  
A;Title: Nucleotide sequence and deduced amino acid sequence of the structural proteins  
A;Reference number: A94346; MUID:87071658; PMID:3024394  
A;Accession: A94346  
A;Molecule type: genomic RNA  
A;Residues: 1-791 <DE1>  
A;Cross-references: UNIPROT:P07564; UNIPARC:UPI00001710BB; GB:M15975  
R;Deubel, V.; Kinney, R.M.; Trent, D.W.  
Virology 165, 234-244, 1988  
A;Title: Nucleotide sequence and deduced amino acid sequence of the nonstructural protein  
A;Reference number: A94378; MUID:88265864; PMID:3388770  
A;Accession: A94378  
A;Molecule type: Genomic RNA  
A;Residues: 792-3391 <DE2>  
A;Cross-references: UNIPARC:UPI0000174A05; GB:M20558  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F;2-114/Product: capsid protein C #status predicted <CPC>  
F;43-59/Domain: transmembrane #status predicted <TM1>  
F;101-117/Domain: transmembrane #status predicted <TM2>  
F;115-280/Product: membrane-associated protein M precursor #status predicted <MPP>  
F;115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
F;206-280/Product: membrane-associated protein M #status predicted <MPM>  
F;268-284/Domain: transmembrane #status predicted <TM3>  
F;281-775/Product: envelope protein E #status predicted <EPE>  
F;727-743/Domain: transmembrane #status predicted <TM4>  
F;757-773/Domain: transmembrane #status predicted <TM5>  
F;776-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
F;1128-1345/Product: nonstructural protein NS2a #status predicted <N2A>  
F;1346-1474/Product: nonstructural protein NS2b #status predicted <N2B>  
F;1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
F;1668-1675/Region: nucleotide-binding motif A (P-loop)  
F;1755-1760/Region: nucleotide-binding motif B  
F;1759-1762/Region: DEAH motif  
F;2094-2243/Product: nonstructural protein NS4a #status predicted <N4A>  
F;2244-2491/Product: nonstructural protein NS5 #status predicted <NS5>  
F;2492-3391/Product: nonstructural protein NS5 (covalent) #status predicted  
F;183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.3%; Score 48; DB 1; Length 3391;  
Best Local Similarity 77.8%; Pred. No. 7.6;  
Matches 7; Conservative 1; Mismatches 0; Gaps 0;





C:Species: dengue virus type 3  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 31-Dec-2004  
C:Accession: A34774  
R:Obatomi, K.; Sumiyoshi, H.  
Virology 176, 643-647, 1990  
A:Title: Complete nucleotide sequence of dengue type 3 virus genome RNA.  
A:Reference number: A34774; MUID:90266483; PMID:2345967  
A:Accession: A34774  
A:Molecule type: genomic RNA  
A:Residues: 1-3390 <OSA>  
A:Cross-references: UNIPROT:P27915; UNIPARC:UPI0000131DFE; GB:M93130; NID:9323468; PIDN:  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F:1-114/Product: capsid protein #status predicted <CAP>  
F:146-67/Domain: transmembrane #status predicted <TM1>  
F:115-280/Product: membrane protein precursor #status predicted <MEP>  
F:115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
F:206-280/Product: membrane protein #status predicted <MEM>  
F:266-280/Domain: transmembrane #status predicted <TM3>  
F:281-773/Product: envelope protein #status predicted <ENV>  
F:724-746/Domain: transmembrane #status predicted <TM4>  
F:753-771/Domain: transmembrane #status predicted <TM5>  
F:774-1184/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1185-1175/Domain: transmembrane #status predicted <TM6>  
F:1185-1343/Product: nonstructural protein NS2a #status predicted <N2A>  
F:1344-1473/Product: nonstructural protein NS2b #status predicted <N2B>  
F:1474-2092/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1667-1674/Region: nucleotide-binding motif A (P-loop)  
F:1754-1759/Region: nucleotide-binding motif B  
F:1759-1761/Region: DEAH motif  
F:2093-2378/Product: nonstructural protein NS4a #status predicted <N4A>  
F:2379-2490/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2491-3390/Product: nonstructural protein NS5 #status predicted <NS>  
F:183,347,433,750,903,980,1132,1188,1661,2300,2304,2386,2456,2702,2712/Binding site: cap

Query Match 85.5%; Score 47; DB 1; Length 3390;  
Best Local Similarity 88.9%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRHP 9  
DB 237 VETWALRHP 245  
RESULT 16  
A2551  
genome polyprotein - dengue virus type 1 (strain Singapore S275/90)  
A:Contains: capsid protein; envelope protein; membrane protein; nonstructural protein NS  
a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: dengue virus type 1  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
C:Accession: A42551  
R:Fu, J.; Tan, B.H.; Yap, E.H.; Chan, Y.C.; Tan, Y.H.  
Virology 188, 953-958, 1992  
A:Title: Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90).  
A:Reference number: A42551; MUID:92263809; PMID:1585663  
A:Accession: A42551  
A:Molecule type: Genomic RNA  
A:Residues: 1-3396 <FUJ>  
A:Cross-references: UNIPROT:P33478; UNIPARC:UPI000002F845; GB:M87512  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F:1-114/Product: capsid protein #status predicted <CAP>  
F:115-281/Product: membrane protein precursor #status predicted <MEP>  
F:115-204/Domain: nonterminal signal sequence #status predicted <SIG>  
F:205-281/Product: membrane protein #status predicted <MEM>  
F:267-279/Domain: transmembrane #status predicted <TM1>  
F:282-774/Product: envelope protein #status predicted <ENV>  
F:753-769/Domain: transmembrane #status predicted <TM2>  
F:775-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1128-1344/Product: nonstructural protein NS2a #status predicted <N2A>  
F:1345-1474/Product: nonstructural protein NS2b #status predicted <N2B>  
F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>

F:1668-1675/Region: nucleotide-binding motif A (P-loop)  
F:1755-1760/Region: nucleotide-binding motif B  
F:1759-1762/Region: DEAH motif  
F:2094-2243/Product: nonstructural protein NS4a #status predicted <N4A>  
F:2244-2492/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2493-3396/Product: nonstructural protein NS5 #status predicted <NS>  
F:183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.5%; Score 47; DB 1; Length 3396;  
Best Local Similarity 88.9%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRHP 9  
DB 237 VETWALRHP 245  
RESULT 17  
GNVVD2  
genome polyprotein - dengue virus type 2 (strain D2-04) (fragment)  
A:Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstruc  
C:Species: dengue virus type 2  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 31-Dec-2004  
C:Accession: JC1007; JC1005  
R:Yang, P.Y.; Lam, S.K.  
Chinese J. Microbiol. Immunol. 11, 341-344, 1991  
A:Title: The nucleotide and encoded amino acid sequences of the structural protein gene (C  
A:Reference number: JC1007  
A:Accession: JC1007  
A:Molecule type: genomic RNA  
A:Residues: 1-775 <YAN>  
A:Cross-references: UNIPROT:P30026; UNIPARC:UPI0000174A06  
A:Note: the authors translated the codons TTA for residue 53 as Phe, AGT for residue 136  
S as Arg, GGC for residue 286 as Ala, and CAG for residue 272 as Leu  
R:Yan, P.Y.; Kautner, I.M.; Koh, C.L.; Lam, S.K.  
Chinese J. Microbiol. Immunol. 11, 9-12, 1991  
A:Title: Nucleotide and encoded amino acid sequences of the nonstructural protein NS1 ge  
A:Reference number: JC1005  
A:Accession: JC1005  
A:Molecule type: genomic RNA  
A:Residues: 776-1127 <YAZ>  
A:Cross-references: UNIPARC:UPI0000174A07  
A:Note: the authors translated the codons GTG for residue 899 as Leu, CTG for residue 952  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; envelope protein; glycoprotein; membrane-associated protein;  
F:1-114/Product: capsid protein C #status predicted <CAP>  
F:101-117/Domain: transmembrane #status predicted <TM1>  
F:115-280/Product: membrane-associated protein M precursor #status predicted <MAM>  
F:115-205/Domain: nonterminal signal sequence #status predicted <SIG>  
F:206-280/Product: membrane-associated protein M #status predicted <MEM>  
F:281-775/Product: envelope protein E #status predicted <ENV>  
F:281-773/Domain: transmembrane #status predicted <TM2>  
F:757-773/Domain: transmembrane #status predicted <TM3>  
F:776-1127/Product: nonstructural protein NS1 #status predicted <NPN>  
F:183,347,433,905,982/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.6%; Score 46; DB 1; Length 1127;  
Best Local Similarity 77.8%; Pred. No. 5.5;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRHP 9  
DB 237 METWILRHP 245  
RESULT 18  
S09223  
membrane protein - dengue virus type 2 (strain M1) (fragment)  
C:Species: dengue virus type 2  
C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 31-Dec-2004  
C:Accession: S09223  
R:Samuel, S.; Koh, C.L.; Pang, T.; Lam, S.K.  
Nucleic Acids Res. 18, 1905, 1990

A;Title: Nucleotide and encoded amino acid sequences of the membrane protein precursor a  
agic fever, dengue shock syndrome or dengue fever.  
A;Reference number: S09223; MUID:90245599; PMID:2336374  
A;Accession: S09223  
A;Molecule type: genomic RNA  
A;Residues: 1-166 <SAM>  
A;Cross-references: UNIPROT:Q67423; UNIPARC:UPI00000F3200; EMBL:X51713; NID:g59309; PIDN  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: membrane protein

Query Match 81.8%; Score 45; DB 2; Length 166;  
Best Local Similarity 66.7%; Pred. No. 1.2;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
: || || || ||  
Db 123 IDTWILRHP 131

RESULT 19  
S09225  
membrane protein - dengue virus type 2 (strain M3) (fragment)  
C;Species: dengue virus type 2  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 31-Dec-2004  
C;Accession: S09225  
R;Samuel, S.; Koh, C.L.; Pang, T.; Lam, S.K.  
Nucleic Acids Res. 18, 1905, 1990  
A;Title: Nucleotide and encoded amino acid sequences of the membrane protein precursor a  
agic fever, dengue shock syndrome or dengue fever.  
A;Reference number: S09223; MUID:90245599; PMID:2336374  
A;Accession: S09225  
A;Molecule type: genomic RNA  
A;Residues: 1-166 <SAM>  
A;Cross-references: UNIPROT:Q67421; UNIPARC:UPI00000E5FFA4; EMBL:X51711; NID:g59305; PIDN  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: membrane protein

Query Match 81.8%; Score 45; DB 2; Length 166;  
Best Local Similarity 66.7%; Pred. No. 1.2;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
: || || || || ||  
Db 123 IQTWILRHP 131

RESULT 20  
PS0043  
genome polyprotein - dengue virus type 2 (strain PU0-218) (fragment)  
N;Contains: envelope protein E; membrane-associated protein M; nonstructural protein NS1  
C;Species: dengue virus type 2  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
C;Accession: PS0043  
R;Gruenberg, A.; Woo, W.S.; Biedrzycka, A.; Wright, P.J.  
J. Gen. Virol. 69, 1391-1398, 1998  
A;Title: Partial nucleotide sequence and deduced amino acid sequence of the structural p  
F;167-661/Product: envelope protein E #status predicted <EP>  
A;Reference number: PS0043; MUID:88258474; PMID:3385407  
A;Accession: PS0043  
A;Molecule type: mRNA  
A;Residues: 1-665 <GRU>  
A;Cross-references: UNIPROT:P18356; UNIPARC:UPI0000178550  
C;Comment: The RNA sequence was obtained from the DDBJ, release 5.0.  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: envelope protein; glycoprotein; membrane protein; nonstructural protein; pol  
F;1-91/Domain: signal sequence #status predicted <SIG>  
F;92-166/Product: membrane-associated protein M #status predicted <MGV>  
F;167-661/Product: envelope protein E #status predicted <EP>  
F;662-665/Product: nonstructural protein NS1 (fragment) #status predicted <NS1>  
F;69,233,319/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 76.4%; Score 42; DB 2; Length 665;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
: || || || || ||  
Db 123 IEIWLHRP 131

RESULT 21  
S09224  
membrane protein - dengue virus type 2 (strain M2) (fragment)  
C;Species: dengue virus type 2  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 31-Dec-2004  
C;Accession: S09224  
R;Samuel, S.; Koh, C.L.; Pang, T.; Lam, S.K.  
Nucleic Acids Res. 18, 1905, 1990  
A;Title: Nucleotide and encoded amino acid sequences of the membrane protein precursor a  
agic fever, dengue shock syndrome or dengue fever.  
A;Reference number: S09223; MUID:90245599; PMID:2336374  
A;Accession: S09224  
A;Molecule type: genomic RNA  
A;Residues: 1-166 <SAM>  
A;Cross-references: UNIPROT:Q67422; UNIPARC:UPI00000F4214; EMBL:X51712; NID:g59307; PIDN  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: membrane protein

Query Match 74.5%; Score 41; DB 2; Length 166;  
Best Local Similarity 66.7%; Pred. No. 5.8;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
: || || || || ||  
Db 123 IVTWILRHP 131

RESULT 22  
AB3184  
probable protein methyltransferase PA3706 [imported] - Pseudomonas aeruginosa (strain PA01)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: AB3184  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brj  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A;Reference number: AB2950; MUID:20437337; PMID:10984043  
A;Accession: AB3184  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-422 <STO>  
A;Cross-references: UNIPROT:Q9HXT5; UNIPARC:UPI00000CSACD; GB:AE004789; GB:AE004091; NID:  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA3706

Query Match 74.5%; Score 41; DB 2; Length 422;  
Best Local Similarity 75.0%; Pred. No. 15;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9  
: || || || || ||  
Db 66 ETWFFRYP 73

RESULT 23  
EB6085  
hypothetical protein yijF [imported] - Escherichia coli (strain O157:H7, substrain EDU93;  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: EB6085  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,  
iller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dialanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: E86085  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-205 <STO>  
A;Cross-references: UNIPROT:Q8X763; UNIPARC:UPI00001659BC; GB:AE005174; NID:g12518859; F  
C;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: yjJF  
C;Superfamily: Escherichia coli hypothetical 23.0K protein b3944  
Query Match 72.7%; Score 40; DB 2; Length 205;  
Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRH 8  
Db 125 LETWFTRH 132  
RESULT 24  
A98238  
hypothetical protein ECs4873 [imported] - Escherichia coli (strain O157:H7, substrain R1  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: A98238  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: A98238  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-205 <HAY>  
A;Cross-references: UNIPROT:Q8X763; UNIPARC:UPI00000DOAB8; GB:BA000007; PIDN:BA838296.1;  
C;Experimental source: strain O157:H7, substrain R1MD 0509952  
C;Genetics:  
A;Gene: ECs4873  
C;Superfamily: Escherichia coli hypothetical 23.0K protein b3944  
Query Match 72.7%; Score 40; DB 2; Length 205;  
Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRH 8  
Db 125 LETWFTRH 132  
RESULT 25  
I78665  
hypothetical 23.0K protein b3944 - Escherichia coli (strain K-12)  
N;Alternate names: hypothetical protein F205  
C;Species: Escherichia coli  
C;Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 09-Jul-2004  
C;Accession: I78665; C65201  
R;Blattner, F.R.; Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.  
Nucleic Acids Res. 21, 5408-5417, 1993  
A;Title: Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89  
A;Reference number: I58303; MUID:94089392; PMID:8265357  
A;Accession: I78665  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-205 <RES>  
A;Cross-references: UNIPROT:P32668; UNIPARC:UPI000013B429; EMBL:U00006; NID:g409785; PID  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: C65201  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA  
A;Residues: 1-205 <BLAT>  
A;Cross-references: UNIPARC:UPI000013B429; GB:AE000468; GB:U00096; NID:gl790374; PIDN:AA  
A;Experimental source: strain K-12, substrain MG1655  
C;Genetics:  
A;Gene: yjJF  
C;Superfamily: Escherichia coli hypothetical 23.0K protein b3944  
Query Match 70.9%; Score 39; DB 2; Length 205;  
Best Local Similarity 75.0%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRH 8  
Db 125 LETWFSRH 132  
RESULT 26  
H95879  
probable sugar ABC transporter permease protein SMB20318 [imported] - Sinorhizobium meli  
C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: H95879  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: H95879  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-343 <KUR>  
A;Cross-references: UNIPROT:Q92WM8; UNIPARC:UPI00000CB4A7; GB:AL591985; PIDN:CAC48704.1;  
A;Experimental source: strain 1021, megaplasmid pSymB  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
peila, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hehault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
C;Genetics:  
A;Gene: SMB20318  
A;Genome: plasmid  
C;Superfamily: l-arabinose transport system permease arah  
Query Match 70.9%; Score 39; DB 2; Length 343;  
Best Local Similarity 83.3%; Pred. No. 27;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 WFLRHP 9  
Db 24 WFLRHP 29  
RESULT 27  
B40098  
colorectal cancer suppressor DCC - rat (fragments)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 27-Mar-1992 #sequence\_revision 10-Apr-1992 #text\_change 24-Jun-1993  
C;Accession: B40098  
R;Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilton  
Science 247, 49-56, 1990  
A;Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.  
A;Reference number: A40098; MUID:90100559; PMID:2294591  
A;Accession: B40098  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-144 <FEA>  
A;Cross-references: UNIPARC:UPI00000422D3; GB:M32287; GB:M32289; GB:M32291  
Query Match 69.1%; Score 38; DB 2; Length 144;



Best Local Similarity 66.7%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
| ||| ||  
Db 53 VPPWFLNHP 61

RESULT 28  
A47666  
structural polyprotein - dengue virus type 4 (fragment)  
N;Contains: capsid protein; envelope glycoprotein; membrane protein precursor  
C;Species: dengue virus type 4  
C;Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 31-Dec-2004  
C;Accession: A47666  
R;Kawano, H.; Rostapshov, V.; Rosen, L.; Lai, C.J.  
J. Virol. 67, 6567-6575, 1993  
A;Title: Genetic determinants of dengue type 4 virus neurovirulence for mice.  
A;Reference number: A47666; PMID:94016840; PMID:8411360  
A;Accession: A47666  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-773 <RAW>  
A;Cross-references: UNIPROT:Q86654; UNIPARC:UPI00000F8175; GB:S66064; NID:G432575; PIDN:  
A;Experimental source: H241-P  
A;Note: sequence extracted from NCBI backbone (NCBIN:138430, NCBIP:138431)  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: glycoprotein; polyprotein

Query Match 69.1%; Score 38; DB 2; Length 773;  
Best Local Similarity 66.7%; Pred. No. 90;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9  
| ||| |||  
Db 236 VESWILRNP 244

RESULT 29  
B96761  
probable protein kinase T9L24.36 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: B96761  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.E.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; PMID:21016719; PMID:11130712  
A;Accession: B96761  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1155 <STO>  
A;Cross-references: UNIPROT:Q9FX38; UNIPARC:UPI000009F5D2; GB:AB005173; NID:g11120796; B  
C;Genetics:  
A;Gene: T9L24.36  
A;Map position: 1

Query Match 69.1%; Score 38; DB 2; Length 1155;  
Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRH 8  
| ||| |  
Db 433 ETWFLAH 439

RESULT 30  
S37034  
DNA-directed DNA polymerase (EC 2.7.7.7) - African swine fever virus  
C;Species: African swine fever virus, ASFV  
C;Date: 09-Dec-1993 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S37034  
R;Martins, A.; Costa, J.V.; Ribeiro, G.  
submitted to the EMBL Data Library, June 1993  
A;Description: Nucleotide sequence of the DNA polymerase gene of African swine fever viru  
A;Reference number: S37034  
A;Accession: S37034  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1244 <MAR>  
A;Cross-references: UNIPROT:P43139; UNIPARC:UPI00001297DA; EMBL:X73330; NID:G397585; PIDN:  
C;Superfamily: African swine fever virus DNA-directed DNA polymerase  
C;Keywords: nucleotidyltransferase

Query Match 69.1%; Score 38; DB 2; Length 1244;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRH 8  
| ||| ||  
Db 561 VEGWFFVRH 568

Search completed: August 31, 2006, 11:51:53  
Job time : 18.25 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 11:33:43 ; Search time 139 Seconds  
(without alignments)  
59.893 Million cell updates/sec

Title: DENGUE\_SEROTYPE1  
Perfect score: 55  
Sequence: 1 vetwflrhp 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Uniprot\_7.2.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	87.3	120	2	Q67424 dengue viru
2	48	87.3	166	2	Q66346 dengue viru
3	48	87.3	280	2	Q8Q264 dengue viru
4	48	87.3	280	2	Q8Q265 dengue viru
5	48	87.3	280	2	Q8Q266 dengue viru
6	48	87.3	280	2	Q8Q267 dengue viru
7	48	87.3	555	1	POLG_DEN2H
8	48	87.3	578	2	Q12290 dengue viru
9	48	87.3	578	2	Q12290 dengue viru
10	48	87.3	661	2	Q3BCV3 dengue viru
11	48	87.3	661	2	Q3BCV4 dengue viru
12	48	87.3	661	2	Q3BCV5 dengue viru
13	48	87.3	661	2	Q3BCX6 dengue viru
14	48	87.3	661	2	Q3BCX7 dengue viru
15	48	87.3	661	2	Q3BCX8 dengue viru
16	48	87.3	661	2	Q3BCX9 dengue viru
17	48	87.3	661	2	Q3BCY0 dengue viru
18	48	87.3	661	2	Q3BCY1 dengue viru
19	48	87.3	661	2	Q3BCY2 dengue viru
20	48	87.3	661	2	Q3BCY3 dengue viru
21	48	87.3	661	2	Q3BCY4 dengue viru
22	48	87.3	661	2	Q3BCY5 dengue viru
23	48	87.3	661	2	Q5Q1B6 dengue viru
24	48	87.3	661	2	Q5V1B7 dengue viru
25	48	87.3	661	2	Q5V1B8 dengue viru
26	48	87.3	661	2	Q5V1B9 dengue viru
27	48	87.3	661	2	Q5V190 dengue viru
28	48	87.3	661	2	Q5V191 dengue viru
29	48	87.3	661	2	Q5V192 dengue viru
30	48	87.3	661	2	Q5V193 dengue viru
31	48	87.3	661	2	Q5V194 dengue viru

ALIGNMENTS

```
RESULT 1
Q67424_9FLAV PRELIMINARY; PRT; 120 AA.
ID Q67424_9FLAV PRELIMINARY; PRT; 120 AA.
AC Q67424;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Genomic RNA for envelope protein E N-term. (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-New Guinea C;
RX MEDLINE=87197230; PubMed=2952760;
RA Biedrzycka A., Cauchi M.R., Bartholomeusz A., Gorman J.J.,
RA Wright P.J.;
RT "Characterization of protease cleavage sites involved in the formation
RT of the envelope glycoprotein and three non-structural proteins of
RT the envelope glycoprotein and three non-structural proteins of
RT J. Gen. Virol. 68:1317-1326(1987).
RL J. Gen. Virol. 68:1317-1326(1987).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; X05375; CA28966.1; -; Genomic_RNA.
DR HSP; Q88653; IOKE.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0019058; P: viral infectious cycle; IEA.
DR InterPro; IPR001599; Flavi_M.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR011998; Vrl GlyE_cen_dim.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Envelope protein.
FT CHAIN 18 92 protein M.
FT NON_TER 93 >120 protein E.
FT NON_TER 1 1
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13329 MW; FF86913787CA5C27 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 120;
Best Local Similarity 77.8%; Pred. No. 2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9
Db 49 IETWILRHP 57
: ||| |||

RESULT 2
Q66346_9FLAV PRELIMINARY; PRT; 166 AA.
ID Q66346_9FLAV PRELIMINARY; PRT; 166 AA.
AC Q66346;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Premembrane polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TH-36;
RA Shiu S.Y.W.;
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; X05375; CA28966.1; -; Genomic_RNA.
DR HSP; Q88653; IOKE.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0019058; P: viral infectious cycle; IEA.
DR InterPro; IPR001599; Flavi_M.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR011998; Vrl GlyE_cen_dim.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Envelope protein.
FT CHAIN 18 92 protein M.
FT NON_TER 93 >120 protein E.
FT NON_TER 1 1
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13329 MW; FF86913787CA5C27 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 120;
Best Local Similarity 77.8%; Pred. No. 2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9
Db 49 IETWILRHP 57
: ||| |||

RESULT 3
Q8QZ64_9FLAV PRELIMINARY; PRT; 280 AA.
ID Q8QZ64_9FLAV PRELIMINARY; PRT; 280 AA.
AC Q8QZ64;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21571640; PubMed=11714970;
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,
RA Holmes E.C., Gould E.A.;
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT for in situ virus evolution and recombination.";
RL J. Gen. Virol. 82:2945-2953(2001).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; AF360863; AAL76291.1; -; Genomic_RNA.
DR SMR; Q8QZ64; 21-100.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0019058; P: viral infectious cycle; IEA.
DR InterPro; IPR001122; Flavi_capsidC.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polypeptide.
FT NON_TER 280 280
SQ SEQUENCE 280 AA; 31847 MW; E889FDD11929CBA7 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 280;
Best Local Similarity 77.8%; Pred. No. 4.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9
Db 237 IETWILRHP 245
: ||| |||
```

```
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; X72849; CAA51363.1; -; mRNA.
DR PIR; S40144; S40144.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019058; P: viral infectious cycle; IEA.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polypeptide.
FT CHAIN 92 >166 membrane protein.
FT NON_TER 1 1
FT NON_TER 166 166
SQ SEQUENCE 166 AA; 18751 MW; F498748D35909639 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 166;
Best Local Similarity 77.8%; Pred. No. 2.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9
Db 123 IETWILRHP 131
: ||| |||

RESULT 3
Q8QZ64_9FLAV PRELIMINARY; PRT; 280 AA.
ID Q8QZ64_9FLAV PRELIMINARY; PRT; 280 AA.
AC Q8QZ64;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21571640; PubMed=11714970;
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,
RA Holmes E.C., Gould E.A.;
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT for in situ virus evolution and recombination.";
RL J. Gen. Virol. 82:2945-2953(2001).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; AF360863; AAL76291.1; -; Genomic_RNA.
DR SMR; Q8QZ64; 21-100.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0019058; P: viral infectious cycle; IEA.
DR InterPro; IPR001122; Flavi_capsidC.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polypeptide.
FT NON_TER 280 280
SQ SEQUENCE 280 AA; 31847 MW; E889FDD11929CBA7 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 280;
Best Local Similarity 77.8%; Pred. No. 4.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9
Db 237 IETWILRHP 245
: ||| |||
```



```
ID POLG DEN2H STANDARD; PRT; 555 AA.
AC P29984;
DT 01-APR-1993, integrated into UniProtKB/Swiss-Prot.
DT 01-APR-1993, sequence version 1.
DT 07-MAR-2006, entry version 43.
DE Genome polyprotein [Contains: Envelope protein M (Matrix protein);
DE Major envelope protein E; Nonstructural protein 1 (NS1)] (Fragment).
OS Dengue virus type 2 (strain TH-36).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=31637;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=92113574; PubMed=1339466;
RA Shiu S.Y.W., Jiang W.R., Porterfield J.S., Gould E.A.;
RT "Envelope protein sequences of dengue virus isolates TH-36 and TH-
RT Sman, and identification of a type-specific genetic marker for dengue
RT and tick-borne flaviviruses."
RL J. Gen. Virol. 73:207-212(1992).
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins
CC (by similarity)
CC -!- MISCELLANEOUS: The virion of this virus is a nucleocapsid covered
CC by a lipoprotein envelope. The envelope contains two proteins: the
CC protein M and glycoprotein E. The nucleocapsid is a complex of
CC protein C and mRNA. In immature particles, there are 60
CC icosahedrally organized trimeric spikes on the surface. Each spike
CC consists of three heterodimers of envelope protein M precursor
CC (pM) and envelope protein E (by similarity).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; D10514; BAA01389.1; -; Genomic_RNA.
DR FIR; JQ1404; JQ1404.
DR HSP; Q88653; IOKE.
DR SMR; P29984; 50-443.
DR InterPro; IPR011999; Flav_glyE_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyE_Ig-like.
DR InterPro; IPR011998; Vrl_glyE_cen_dm.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycop_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Capsid protein; Core protein; Envelope protein; Glycoprotein;
KW Membrane; Polyprotein; Structural protein; Transmembrane.
FT CHAIN <1 49
FT /FTID=PRO_0000037922.
FT Major envelope protein E.
FT /FTID=PRO_0000037923.
FT Nonstructural protein 1.
FT /FTID=PRO_0000037924.
FT Potential.
FT TRANSMEM 37 53
FT TRANSMEM 496 512
FT TRANSMEM 526 542
FT CARBOHYD 116 116
FT CARBOHYD 202 202
FT DISULFID 52 79
FT DISULFID 109 170
FT DISULFID 123 154
FT DISULFID 141 165
FT DISULFID 234 334
FT DISULFID 351 382
FT NON_TER 1
FT NON_TER 555 555
SQ SEQUENCE 555 AA; 61243 MW; F8DEA740BB4D8DF CRC64;
Query Match 87.3%; Score 48; DB 1; Length 555;
Best Local Similarity 77.8%; Pred. No. 9.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 VETWFLRHP 9
Db 6 IETWILRHP 14
```

```
RESULT 8
Q3ZPU0_9FLAV
ID Q3ZPU0_9FLAV PRELIMINARY; PRT; 565 AA.
AC Q3ZPU0;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3930;
RA Pyke A.T., Hanna J., Richards A., Taylor C.T., Morgan A.,
RA Humphreys J., Brookes D., Smith G.A.;
RT "Defining Dengue in the New Millennium.";
RL Arbovirus Res. Aust. 0:0-0(2005).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; AY06011; AAW62469.1; -; Genomic_RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyE_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR001157; Flavi_NSI.
DR InterPro; IPR000336; Flv_glyE_Ig-like.
DR InterPro; IPR011998; Vrl_glyE_cen_dm.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycop_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF00948; Flavi_NSI; 1.
DR ProDom; PD001496; Flavi_NSI; 1.
DR Polyprotein.
FT NON_TER 1
FT NON_TER 565 565
SQ SEQUENCE 565 AA; 61930 MW; 17DC94EEC53B3EF6 CRC64;
Query Match 87.3%; Score 48; DB 2; Length 565;
Best Local Similarity 77.8%; Pred. No. 9.5;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 VETWFLRHP 9
Db 2 IETWILRHP 10
```

```
RESULT 9
O12290_9FLAV
ID O12290_9FLAV PRELIMINARY; PRT; 578 AA.
AC O12290;
DT 01-JUL-1997, integrated into UniProtKB/TrEMBL.
DT 01-JUL-1997, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-Torres Strait 1;
RA Serafin I.L., Phillips D.A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC
```

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----

DR EMBL: AF004019; AAB61366.1; -; mRNA.  
DR HSSP: Q88653; IOKE.  
DR SMR: Q12290; 48-441.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0019028; C: viral capsid; IEA.  
DR GO: GO:0019031; C: viral envelope; IEA.  
DR GO: GO:0005198; F: structural molecule activity; IEA.  
DR GO: GO:0019058; P: viral infectious cycle; IEA.  
DR InterPro: IPR011999; Flav\_glyc\_cen\_dm.  
DR InterPro: IPR000069; Flavi\_M.  
DR InterPro: IPR001157; Flavi\_NSI.  
DR InterPro: IPR000336; Flv\_glyc\_ig-like.  
DR InterPro: IPR011998; Vrl\_glyc\_cen\_dm.  
DR Pfam: PF02832; Flavi\_glycop\_C; 1.  
DR Pfam: PF00869; Flavi\_glycoprot; 1.  
DR Pfam: PF01004; Flavi\_M; 1.  
DR Pfam: PF00948; Flavi\_NSI; 1.  
DR ProDom: PD001496; Flavi\_NSI; 1.  
KW Polypeptide.  
FT CHAIN <1 47 membrane protein.  
FT CHAIN 48 542 envelope protein.  
FT CHAIN 543 >578 nonstructural protein 1.  
FT NON\_TER 1 1  
FT NON\_TER 578 578  
SQ SEQUENCE 578 AA; 63606 MW; 1C03A7CFD72C567D CRC64;

Query Match 87.3%; Score 48; DB 2; Length 578;  
Best Local Similarity 77.8%; Pred. No. 9.7;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9  
:|||||  
Db 4 IETWILRHP 12

RESULT 10  
ID Q3BCV3\_9FLAV PRELIMINARY; PRT; 661 AA.  
AC Q3BCV3;  
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.  
DT 22-NOV-2005, sequence version 1.  
DT 07-FEB-2006, entry version 3.  
DE Polypeptide (Fragment).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; Dengue virus group.  
OX NCBI\_TaxID=11060;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=01st193/2001;  
RX PubMed=16222028;  
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N., Morita K.;  
RA "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES: GENOTYPE SHIFT AND LOCAL EVOLUTION.";  
RT Am. J. Trop. Med. Hyg. 73:796-802(2005).  
RL [2]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=01st193/2001;  
RX PubMed=16222028;  
RA Salda L.T.;  
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----

EMBL: AY786398; AAX18216.1; -; Genomic\_RNA.  
GO: GO:0016021; C: integral to membrane; IEA.  
GO: GO:0019028; C: viral capsid; IEA.  
GO: GO:0019031; C: viral envelope; IEA.  
GO: GO:0005198; F: structural molecule activity; IEA.

DR GO: GO:0019058; P: viral infectious cycle; IEA.  
KW Polypeptide.  
FT NON\_TER 1 1  
FT NON\_TER 661 661  
SQ SEQUENCE 661 AA; 73049 MW; 2A644DEADA728CF2 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9  
:|||||  
Db 123 IETWILRHP 131

RESULT 11  
ID Q3BCV4\_9FLAV PRELIMINARY; PRT; 661 AA.  
AC Q3BCV4;  
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.  
DT 22-NOV-2005, sequence version 1.  
DT 07-FEB-2006, entry version 3.  
DE Polypeptide (Fragment).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; Dengue virus group.  
OX NCBI\_TaxID=11060;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=0003/2000/human;  
RX PubMed=16222028;  
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N., Morita K.;  
RA "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES: GENOTYPE SHIFT AND LOCAL EVOLUTION.";  
RT Am. J. Trop. Med. Hyg. 73:796-802(2005).  
RL [2]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=0003/2000/human;  
RX Salda L.T.D.;  
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----

EMBL: AY786397; AAX18215.1; -; Genomic\_RNA.  
GO: GO:0016021; C: integral to membrane; IEA.  
GO: GO:0019028; C: viral capsid; IEA.  
GO: GO:0019031; C: viral envelope; IEA.  
GO: GO:0005198; F: structural molecule activity; IEA.  
GO: GO:0019058; P: viral infectious cycle; IEA.  
KW Polypeptide.  
FT NON\_TER 1 1  
FT NON\_TER 661 661  
SQ SEQUENCE 661 AA; 73072 MW; 654A28D6B96FBSA8 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9  
:|||||  
Db 123 IETWILRHP 131

RESULT 12  
ID Q3BCV5\_9FLAV PRELIMINARY; PRT; 661 AA.  
AC Q3BCV5;  
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.  
DT 22-NOV-2005, sequence version 1.  
DT 07-FEB-2006, entry version 3.  
DE Polypeptide (Fragment).



```

DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99Sa695/1999;
RX PubMed=16222028;
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99Sa695/1999;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786373; AAX18191.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polypeptide.
FT NON_TER 1 661
FT SEQUENCE 661 AA; 73086 MW; 899A28D6B96FB5B0 CRC64;
SQ
Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 16
Q3BCX9_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX9;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCMC60/1998;
RX PubMed=16222028;
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCMC60/1998;
RX PubMed=16222028;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786373; AAX18191.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polypeptide.
FT NON_TER 1 661
FT SEQUENCE 661 AA; 73086 MW; 899A28D6B96FB5B0 CRC64;
SQ
Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 17
Q3BCY0_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY0;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C127/1998;
RX PubMed=16222028;
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C127/1998;
RX Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786371; AAX18189.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polypeptide.
FT NON_TER 1 661
FT SEQUENCE 661 AA; 73166 MW; 84C50AFD2358F08C CRC64;
SQ
Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VETWFLRHP 9
Db 123 IETWILRHP 131

```

```

CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786372; AAX18190.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polypeptide.
FT NON_TER 1 661
FT SEQUENCE 661 AA; 73072 MW; 654A28D6B96FB5A8 CRC64;
SQ
Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 17
Q3BCY0_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY0;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C127/1998;
RX PubMed=16222028;
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C127/1998;
RX Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786371; AAX18189.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polypeptide.
FT NON_TER 1 661
FT SEQUENCE 661 AA; 73166 MW; 84C50AFD2358F08C CRC64;
SQ
Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VETWFLRHP 9
Db 123 IETWILRHP 131

```

```

RESULT 18
Q3BCY1_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY1;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CSMC7/1996;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CSMC7/1996;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; AY786370; AAX18188.1; -; Genomic RNA.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0019031; C:viral envelope; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73054 MW; 75134447E73C46F CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 19
Q3BCY2_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY2;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRL3/1996;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRL3/1996;
RX PubMed=16222028;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; AY786368; AAX18186.1; -; Genomic RNA.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0019031; C:viral envelope; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73096 MW; CF865AAE54ADE0F1 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 20
Q3BCY3_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY3;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH97/1995;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH97/1995;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; AY786368; AAX18186.1; -; Genomic RNA.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0019031; C:viral envelope; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73096 MW; CF865AAE54ADE0F1 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db 123 IETWILRHP 131
```



```

RT  GENOTYPE SHIFT AND LOCAL EVOLUTION." ;
RL  Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=SLMC125/1995;
RA  Salda L.T.D.;
RL  Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
DR  EMBL; AY786366; AX18184.1; -; Genomic RNA.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0019028; C:viral capsid; IEA.
DR  GO; GO:0019031; C:viral envelope; IEA.
DR  GO; GO:0005198; F:structural molecule activity; IEA.
DR  GO; GO:0019058; P:viral infectious cycle; IEA.
KW  Polyprotein.
FT  NON_TER 1
FT  NON_TER 661
FT  NON_TER 661
SQ  SEQUENCE 661 AA; 73072 MW; 654A28D6B96FB5A8 CRC64;

    Query Match      87.3%; Score 48; DB 2; Length 661;
    Best Local Similarity 77.8%; Pred. No. 11;
    Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  1 VETWFLRHP 9
DB  123 IETWILRHP 131
    :||| |||
    :||| |||

RESULT 23
Q5QIB6 9FLAV PRELIMINARY; PRT; 661 AA.
ID  Q5QIB6 9FLAV PRELIMINARY; PRT; 661 AA.
AC  Q5QIB6;
DT  04-JAN-2005, integrated into UniProtKB/TrEMBL.
DT  04-JAN-2005, sequence version 1.
DT  07-FEB-2006, entry version 5.
DE  Polyprotein (Fragment).
DE  Polypeptide (Fragment).
OS  Dengue virus type 2.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Flavivirus; Dengue virus group.
NCBI_TaxID=11060;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=BC134-Merida-94;
RX  PubMed=15516647;
RA  Llorca-Ojeda M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA  Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA  Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA  Beaty B.J.;
RA  "Introduction of the American/Asian genotype of dengue 2 virus into
RT  the Yucatan State of Mexico.";
RL  Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC  -----
CC  Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC  Distributed under the Creative Commons Attribution-NoDerivs License
CC  -----
DR  EMBL; AY466449; AAS45234.1; -; mRNA.
DR  SMR; Q5QIB6; 167-560.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0019028; C:viral capsid; IEA.
DR  GO; GO:0019031; C:viral envelope; IEA.
DR  GO; GO:0005198; F:structural molecule activity; IEA.
DR  GO; GO:0019058; P:viral infectious cycle; IEA.
DR  InterPro; IPR011999; Flavi glye_cen_dm.
DR  InterPro; IPR000069; Flavi_M.
DR  InterPro; IPR002535; Flavi_M.
DR  InterPro; IPR000336; Flv glye_ig-like.
DR  InterPro; IPR011998; Vri glye_cen_dim.
DR  Pfam; PF02832; Flavi glycop C; 1.
DR  Pfam; PF00869; Flavi glycoprot; 1.
DR  Pfam; PF01004; Flavi_M; 1.

```

```
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73207 MW; A919612986E04157 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 24
QSVI87_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI87;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN 1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=13382/Tizimin 02;
RX PubMed=15516647;
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL; AY449684; AAS14975.1; -; Genomic_RNA.
SMR; QSVI87; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000699; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73119 MW; CE2051C17F40A623 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 25
QSVI87_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI87;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN 1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=13382/Tizimin 02;
RX PubMed=15516647;
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL; AY449684; AAS14975.1; -; Genomic_RNA.
SMR; QSVI87; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000699; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73119 MW; CE2051C17F40A623 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 26
QSVI89_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI89;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN 1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=12021/Oxkutzab 01;
RX PubMed=15516647;
RA Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
```

```
QSVI88_9FLAV PRELIMINARY; PRT; 661 AA.
ID QSVI88_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI88;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN 1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=13381/Chochola 02;
RX PubMed=15516647;
RA Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL; AY449683; AAS14974.1; -; Genomic_RNA.
SMR; QSVI88; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000699; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73092 MW; 482C14A6B3B179FA CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 26
QSVI89_9FLAV PRELIMINARY; PRT; 661 AA.
ID QSVI89_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI89;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN 1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=12021/Oxkutzab 01;
RX PubMed=15516647;
RA Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
```

RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,  
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,  
RA "Introduction of the American/Asian genotype of dengue 2 virus into  
RT the Yucatan State of Mexico";  
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AY449682; AAS14973.1; -; Genomic\_RNA.  
DR SMR; QSV190; 167-560.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0019058; P:viral infectious cycle; IEA.  
DR InterPro; IPR011999; Flav\_glyc\_cen\_dm.  
DR InterPro; IPR000069; Flavi\_M.  
DR InterPro; IPR002535; Flavi\_propep.  
DR InterPro; IPR000336; Flv\_glyc\_ig-like.  
DR Pfam; PF02832; Flavi\_glycop\_C; 1.  
DR Pfam; PF00869; Flavi\_glycoprot; 1.  
DR Pfam; PF01004; Flavi\_M; 1.  
DR Pfam; PF01570; Flavi\_propep; 1.  
DR Polyprotein.  
KW POLYPROTEIN.  
FT NON\_TER 1  
FT NON\_TER 661  
SQ SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;  
Query Match 87.3%; Score 48; DB 2; Length 661;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRHP 9  
DB 123 IETWILRHP 131  
:|||||  
RESULT 27  
QSV190\_9FLAV PRELIMINARY; PRT; 661 AA.  
AC QSV190;  
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 07-DEC-2004, sequence version 1.  
DT 07-FEB-2006, entry version 8.  
DE Polyprotein (Fragment).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; Dengue virus group.  
OX NCBI\_TaxID=11060;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=1936/St. Elena 01;  
RX PubMed=15516647;  
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,  
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,  
RA "Introduction of the American/Asian genotype of dengue 2 virus into  
RT the Yucatan State of Mexico";  
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AY449681; AAS14972.1; -; Genomic\_RNA.  
DR SMR; QSV190; 167-560.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.

DR GO; GO:0019058; P:viral infectious cycle; IEA.  
DR InterPro; IPR011999; Flav\_glyc\_cen\_dm.  
DR InterPro; IPR000069; Flavi\_M.  
DR InterPro; IPR002535; Flavi\_propep.  
DR InterPro; IPR000336; Flv\_glyc\_ig-like.  
DR Pfam; PF02832; Flavi\_glycop\_C; 1.  
DR Pfam; PF00869; Flavi\_glycoprot; 1.  
DR Pfam; PF01004; Flavi\_M; 1.  
DR Pfam; PF01570; Flavi\_propep; 1.  
DR Polyprotein.  
KW POLYPROTEIN.  
FT NON\_TER 1  
FT NON\_TER 661  
SQ SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;  
Query Match 87.3%; Score 48; DB 2; Length 661;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRHP 9  
DB 123 IETWILRHP 131  
:|||||  
RESULT 28  
QSV191\_9FLAV PRELIMINARY; PRT; 661 AA.  
AC QSV191;  
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 07-DEC-2004, sequence version 1.  
DT 07-FEB-2006, entry version 8.  
DE Polyprotein (Fragment).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; Dengue virus group.  
OX NCBI\_TaxID=11060;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=12914/Tekax 01;  
RX PubMed=15516647;  
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,  
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,  
RA "Introduction of the American/Asian genotype of dengue 2 virus into  
RT the Yucatan State of Mexico";  
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AY449680; AAS14971.1; -; Genomic\_RNA.  
DR SMR; QSV191; 167-560.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0019058; P:viral infectious cycle; IEA.  
DR InterPro; IPR011999; Flav\_glyc\_cen\_dm.  
DR InterPro; IPR000069; Flavi\_M.  
DR InterPro; IPR002535; Flavi\_propep.  
DR InterPro; IPR000336; Flv\_glyc\_ig-like.  
DR InterPro; IPR011998; Flv\_glyc\_ig-like.  
DR Pfam; PF02832; Flavi\_glycop\_C; 1.  
DR Pfam; PF00869; Flavi\_glycoprot; 1.  
DR Pfam; PF01004; Flavi\_M; 1.  
DR Pfam; PF01570; Flavi\_propep; 1.  
DR Polyprotein.  
KW POLYPROTEIN.  
FT NON\_TER 1  
FT NON\_TER 661  
SQ SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;  
Query Match 87.3%; Score 48; DB 2; Length 661;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VETWFLRHP 9  
DB 123 IETWILRHP 131  
:|||||  
RESULT 28  
QSV191\_9FLAV PRELIMINARY; PRT; 661 AA.  
AC QSV191;  
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 07-DEC-2004, sequence version 1.  
DT 07-FEB-2006, entry version 8.  
DE Polyprotein (Fragment).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; Dengue virus group.  
OX NCBI\_TaxID=11060;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=12914/Tekax 01;  
RX PubMed=15516647;  
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,  
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,  
RA "Introduction of the American/Asian genotype of dengue 2 virus into  
RT the Yucatan State of Mexico";  
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AY449680; AAS14971.1; -; Genomic\_RNA.  
DR SMR; QSV191; 167-560.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0019058; P:viral infectious cycle; IEA.  
DR InterPro; IPR011999; Flav\_glyc\_cen\_dm.  
DR InterPro; IPR000069; Flavi\_M.  
DR InterPro; IPR002535; Flavi\_propep.  
DR InterPro; IPR000336; Flv\_glyc\_ig-like.  
DR InterPro; IPR011998; Flv\_glyc\_ig-like.  
DR Pfam; PF02832; Flavi\_glycop\_C; 1.  
DR Pfam; PF00869; Flavi\_glycoprot; 1.  
DR Pfam; PF01004; Flavi\_M; 1.  
DR Pfam; PF01570; Flavi\_propep; 1.  
DR Polyprotein.  
KW POLYPROTEIN.  
FT NON\_TER 1  
FT NON\_TER 661  
SQ SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;  
Query Match 87.3%; Score 48; DB 2; Length 661;

Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9  
Db 123 IETWILRHP 131

RESULT 29  
ID QSVI92\_9FLAV PRELIMINARY; PRT; 661 AA.  
AC QSVI92;  
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 07-DEC-2004, sequence version 1.  
DE Polypeptide (Fragment).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; Dengue virus group.  
OX NCBI\_TaxID=11060;

NUCLEOTIDE SEQUENCE.  
STRAIN=C-932/Chilpancingo 97;  
PubMed=15516647;  
Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,  
Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,  
Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,  
Beatty B.J.;  
"Introduction of the American/Asian genotype of dengue 2 virus into  
the Yucatan State of Mexico";  
Am. J. Trop. Med. Hyg. 71:485-492 (2004).

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AY449679; AAS14970.1; -; Genomic\_RNA.  
SMR; QSVI92; 167-560.  
GO; GO:0016021; C:integral to membrane; IEA.  
GO; GO:0019028; C:viral capsid; IEA.  
GO; GO:0019031; C:viral envelope; IEA.  
GO; GO:0005198; F:structural molecule activity; IEA.  
GO; GO:0019058; P:viral infectious cycle; IEA.  
InterPro; IPR011999; Flav\_glyc\_cen\_dm.  
InterPro; IPR000069; Flavi\_M.  
InterPro; IPR002535; Flavi\_propep.  
InterPro; IPR000336; Flv\_glyc\_ig-like.  
Pfam; PF02832; Flavi\_glycop\_C; 1.  
Pfam; PF00869; Flavi\_glycoprot; 1.  
Pfam; PF01004; Flavi\_M; 1.  
Pfam; PF01570; Flavi\_propep; 1.  
KW Polyprotein.  
FT NON\_TER 1  
FT NON\_TER 661  
SQ SEQUENCE 661 AA; 73024 MW; 367D6F1A9F25932B CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9  
Db 123 IETWILRHP 131

RESULT 30  
ID QSVI93\_9FLAV PRELIMINARY; PRT; 661 AA.  
AC QSVI93;  
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 07-DEC-2004, sequence version 1.  
DE Polypeptide (Fragment).

Dengue virus type 2.  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Flavivirus; Dengue virus group.  
NCBI\_TaxID=11060;  
[1]  
NUCLEOTIDE SEQUENCE.  
STRAIN=C-932/Acapulco 97;  
PubMed=15516647;  
Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,  
Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,  
Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,  
Beatty B.J.;  
"Introduction of the American/Asian genotype of dengue 2 virus into  
the Yucatan State of Mexico";  
Am. J. Trop. Med. Hyg. 71:485-492 (2004).

Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
Distributed under the Creative Commons Attribution-NoDerivs License

EMBL; AY449678; AAS14969.1; -; Genomic\_RNA.  
SMR; QSVI93; 167-560.  
GO; GO:0016021; C:integral to membrane; IEA.  
GO; GO:0019028; C:viral capsid; IEA.  
GO; GO:0019031; C:viral envelope; IEA.  
GO; GO:0005198; F:structural molecule activity; IEA.  
GO; GO:0019058; P:viral infectious cycle; IEA.  
InterPro; IPR011999; Flav\_glyc\_cen\_dm.  
InterPro; IPR000069; Flavi\_M.  
InterPro; IPR002535; Flavi\_propep.  
InterPro; IPR000336; Flv\_glyc\_ig-like.  
Pfam; PF02832; Flavi\_glycop\_C; 1.  
Pfam; PF00869; Flavi\_glycoprot; 1.  
Pfam; PF01004; Flavi\_M; 1.  
Pfam; PF01570; Flavi\_propep; 1.  
KW Polyprotein.  
FT NON\_TER 1  
FT NON\_TER 661  
SQ SEQUENCE 661 AA; 73024 MW; 0E74A2AC438791A1 CRC64;

Query Match 87.3%; Score 48; DB 2; Length 661;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VETWFLRHP 9  
Db 123 IETWILRHP 131

Search completed: August 31, 2006, 11:43:09  
Job time : 139 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 11:33:43 ; Search time 110.25 Seconds  
(without alignments)  
37.324 Million cell updates/sec

Title: DENGUE\_SEROTYPE2

Perfect score: 55

Sequence: 1 ietwflrhp 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_8.\*

1: geneseqp1980s.\*

2: geneseqp1990s.\*

3: geneseqp2000s.\*

4: geneseqp2001s.\*

5: geneseqp2002s.\*

6: geneseqp2003as.\*

7: geneseqp2003bs.\*

8: geneseqp2004s.\*

9: geneseqp2005s.\*

10: geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	39	9	Adw12582 M1-40/DEN
2	55	100.0	48	9	Adw12588 p(95-114)
3	49	89.1	9	9	Adw12595 M32-40/DE
4	49	89.1	21	9	Adw12594 M20-40/DE
5	49	89.1	32	9	Adw12593 M10-40/DEN
6	49	89.1	39	9	Adw12576 M1-40/DEN
7	49	89.1	40	5	Aae17432 Dengue (D
8	49	89.1	40	5	Adw12578 M1-40/YF
9	49	89.1	40	5	Aae17433 (95-114)E
10	49	89.1	167	8	Adn37497 Dengue vi
11	49	89.1	171	8	Adn37493 Dengue vi
12	49	89.1	171	8	Adn37496 Dengue vi
13	49	89.1	635	2	Aaw75410 Fusion pr
14	49	89.1	675	8	Adn37628 Dengue vi
15	49	89.1	675	8	Adn37518 Dengue vi
16	49	89.1	675	8	Adn37612 Dengue vi
17	49	89.1	675	8	Adn37626 Dengue vi
18	49	89.1	677	2	Aaw75411 Fusion pr
19	49	89.1	677	8	Adn37613 Dengue vi
20	49	89.1	681	8	Adn37603 Dengue vi
21	49	89.1	681	8	Adn37517 Dengue vi
22	49	89.1	685	6	Abp57874 Plasmid p
23	49	89.1	685	6	Abp57876 Plasmid p

24	89.1	685	6	ABP57875	Abp57875 Plasmid p
25	49	89.1	1127	2	Aaw09409 Dengue vi
26	49	89.1	1127	2	Aay05522 Dengue vi
27	49	89.1	1127	7	Adl98086 Dengue vi
28	49	89.1	1127	8	Adq28716 Dengue vi
29	49	89.1	3388	6	AAE35314 Dengue vi
30	49	89.1	3391	2	AAW13166 Proteins
31	49	89.1	3391	2	AAW06591 Polyprote
32	49	89.1	3391	2	AAW06590 Polyprote
33	49	89.1	3391	4	AAE07987 Attenuate
34	49	89.1	3391	4	AAE07986 Wild-type
35	49	89.1	3391	8	ADG93314 DEN2 (Ton
36	47	85.5	9	ADM12597	M32-40/DE
37	46	83.6	40	5	AAE17431 Dengue (D
38	46	83.6	48	5	AAE17437 (95-114)E
39	46	83.6	55	5	AAE17438 p(95-114)
40	46	83.6	167	8	Adn37494 Dengue vi
41	46	83.6	167	8	Adn37501 Dengue vi
42	46	83.6	167	8	Adn37498 Dengue vi
43	46	83.6	167	8	Adn37492 Dengue vi
44	46	83.6	167	8	Adn37500 Dengue vi
45	46	83.6	675	8	Adn37624 Dengue vi
46	46	83.6	675	8	Adn37519 Dengue vi
47	46	83.6	675	8	Adn37521 Dengue vi
48	46	83.6	675	8	Adn37616 Dengue vi
49	46	83.6	675	8	Adn37523 Dengue vi
50	46	83.6	675	8	Adn37621 Dengue vi
51	46	83.6	675	8	Adn37604 Dengue vi
52	46	83.6	675	8	Adn37614 Dengue vi
53	46	83.6	675	8	Adn37618 Dengue vi
54	46	83.6	675	8	Adn37620 Dengue vi
55	46	83.6	675	8	Adn37615 Dengue vi
56	46	83.6	675	8	Adn37611 Dengue vi
57	46	83.6	676	8	Adn37619 Dengue vi
58	46	83.6	677	8	Adn37617 Dengue vi
59	46	83.6	677	8	Adn37522 DEN-1/DEN
60	46	83.6	677	8	Adn37602 Dengue vi
61	46	83.6	677	8	Adn37515 Dengue vi
62	46	83.6	679	8	AdS76179 Heterodim
63	46	83.6	681	8	Adn37622 Dengue vi
64	46	83.6	684	8	ADR87180 Dengue vi
65	46	83.6	715	2	AAW06593 Amino aci
66	46	83.6	774	8	ADG93320 DEN1 (Pue
67	46	83.6	775	8	ADG93318 DEN1 (Pue
68	46	83.6	798	2	AAW06592 Amino aci
69	46	83.6	3389	4	AAE07984 Dengue vi
70	46	83.6	3390	4	AAE07989 Wild-type
71	46	83.6	3390	4	AAE07990 Attenuate
72	46	83.6	3391	4	AAE07982 Dengue vi
73	46	83.6	3391	4	AAE07983 Dengue vi
74	46	83.6	3391	4	AAE07993 Attenuate
75	46	83.6	3392	4	AAE07981 Wild-type
76	46	83.6	3392	4	AAE07980 DEN1-S275
77	46	83.6	3396	2	AAR43662 M1-40/DEN
78	45	81.8	39	9	Adw12599 PUO-218 s
79	43	78.2	150	1	AAp91166 Dengue-2
80	43	78.2	661	4	AAh84901 Dengue vi
81	43	78.2	661	9	AED66282 Dengue vi
82	42	76.4	27	8	Adn11192 Peptide m
83	42	76.4	27	8	Adn11216 Peptide m
84	42	76.4	278	8	ADQ25888 Human GPC
85	42	76.4	611	6	ABU22889 Protein e
86	42	76.4	826	5	ABb07253 Human nov
87	42	76.4	827	6	ABU07568 Human sec
88	42	76.4	904	4	ABG09947 Novel hum
89	42	76.4	924	5	AAb71323 Human GCR
90	42	76.4	953	7	AAE34415 Human nov
91	42	76.4	994	5	ABb07252 Human nov
92	42	76.4	994	5	AAU9808 Novel hum
93	42	76.4	994	7	ADE34425 Human G-p
94	42	76.4	994	8	ADO28977 Human nov
95	42	76.4	994	8	ADQ25892 Human.gua
96	42	76.4	1018	5	AAE25061 Human G-p

97 42 76.4 1070 6 ABU07567 Human sec  
 98 42 76.4 1131 4 ABG11655  
 99 42 76.4 1232 7 ADF70474  
 100 42 76.4 3390 8 ADG93316

## ALIGNMENTS

## RESULT 1

ADW12582  
 ID ADW12582 standard; peptide; 39 AA.  
 AC ADW12582;  
 XX  
 DT 24-MAR-2005 (first entry)  
 XX  
 DE M1-40/DEN-2 (F36) mutant protein.  
 XX  
 KW Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 KW DEN; dengue; mutant; mutein.  
 XX  
 OS Dengue virus.  
 XX  
 FN US2004266987-A1.  
 XX  
 PD 30-DEC-2004.  
 XX  
 XX 30-JUN-2003; 2003US-00608029.  
 PF  
 XX 30-JUN-2003; 2003US-00608029.  
 PR  
 XX (INSP ) INST PASTEUR.  
 PA  
 XX Despres P, Catteau A;  
 PI  
 XX WPI; 2005-047647/05.  
 DR  
 XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
 XX as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.  
 XX  
 PS Example 1; SEQ ID NO 29; 30pp; English.  
 XX  
 CC The present invention relates to an isolated and purified ApoptoM  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC M1-40/DEN (dengue)-2 (F36) mutant protein.  
 XX  
 SQ Sequence 39 AA;

Query Match 100.0%; Score 55; DB 9; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 0.018;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9

Db 31 IETWFLRHP 39

## RESULT 2

ADW12588  
 ID ADW12588 standard; protein; 48 AA.  
 XX  
 AC ADW12588;  
 XX  
 DT 24-MAR-2005 (first entry)  
 XX

DE p(95-114) EGFP(M1-M40)DEN-2 (136F) plasmid DNA encoded protein #3.  
 XX  
 KW Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 KW DEN; dengue; EGFP; enhanced green fluorescent protein.  
 XX  
 OS Dengue virus.  
 OS Chimeric.  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2 /note= "Encoded by GGC"  
 FT Misc-difference 4 /note= "Encoded by GAC"  
 FT Misc-difference 13.44  
 FT /note= "Encoded by GTTTC"

XX US2004266987-A1.

XX 30-DEC-2004.

XX 30-JUN-2003; 2003US-00608029.

XX 30-JUN-2003; 2003US-00608029.

XX (INSP ) INST PASTEUR.

XX Despres P, Catteau A;

XX WPI; 2005-047647/05.

DR N-PSDB; ADW12589.

XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
 PT as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.

XX Disclosure; SEQ ID NO 35; 30pp; English.

XX The present invention relates to an isolated and purified ApoptoM  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC p(95-114) EGFP (enhanced green fluorescent protein) (M1-M40)DEN (dengue)-2  
 CC (136F) plasmid DNA encoded protein.

XX Sequence 48 AA;

Query Match 100.0%; Score 55; DB 9; Length 48;

Best Local Similarity 100.0%; Pred. No. 0.022;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9

Db 40 IETWFLRHP 48

## RESULT 3

ADW12595  
 ID ADW12595 standard; peptide; 9 AA.  
 XX  
 AC ADW12595;  
 XX  
 DT 24-MAR-2005 (first entry)  
 XX  
 DE M32-40/DEN-2 mutant protein #1.

XX Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 KW DEN; dengue; mutant; mutein.

```

XX OS Dengue virus.
XX PN US2004266987-A1.
XX PD 30-DEC-2004.
XX PF 30-JUN-2003; 2003US-00608029.
XX PR 30-JUN-2003; 2003US-00608029.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX DR WPI; 2005-047647/05.
XX PT New isolated and purified Apoptom peptide comprises 9 amino acids, useful
PT as a vaccine for preventing or treating pathological conditions from non-
PT specific febrile illnesses to severe hemorrhagic manifestations or
PT encephalitic syndromes.
XX PS Example 3; Fig 4; 30pp; English.
XX CC The present invention relates to an isolated and purified Apoptom
CC peptide. The invention is useful as a vaccine for the prevention and
CC treatment of pathological conditions from non-specific febrile illnesses
CC to severe hemorrhagic manifestations, encephalitic syndromes and these
CC pathological conditions are linked to Flavivirus infection or cancers.
CC The invention is also useful in gene therapy. The present sequence is a
CC M32-40/DEN (dengue)-2 mutant protein.
XX SQ Sequence 9 AA;

Query Match 89.1%; Score 49; DB 9; Length 9;
Best Local Similarity 88.9%; Pred. No. 2.1e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 IETWFLRHP 9
Db 1 IETWILRHP 9

RESULT 4
ADW12594
ID ADW12594 standard; peptide; 21 AA.
XX AC ADW12594;
XX DT 24-MAR-2005 (first entry)
XX DE M20-40/DEN-2 mutant protein.
XX KW Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX KW DEN; dengue; mutant; mutein.
XX OS Dengue virus.
XX OS US2004266987-A1.
XX PN 30-DEC-2004.
XX PF 30-JUN-2003; 2003US-00608029.
XX PR 30-JUN-2003; 2003US-00608029.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX DR WPI; 2005-047647/05.
XX PT New isolated and purified Apoptom peptide comprises 9 amino acids, useful
XX as a vaccine for preventing or treating pathological conditions from non-
XX specific febrile illnesses to severe hemorrhagic manifestations or
XX encephalitic syndromes.
XX PS Example 3; Fig 4; 30pp; English.
XX CC The present invention relates to an isolated and purified Apoptom
XX peptide. The invention is useful as a vaccine for the prevention and
XX treatment of pathological conditions from non-specific febrile illnesses
XX to severe hemorrhagic manifestations, encephalitic syndromes and these
XX pathological conditions are linked to Flavivirus infection or cancers.
XX The invention is also useful in gene therapy. The present sequence is a
XX M32-40/DEN (dengue)-2 mutant protein.
XX SQ Sequence 9 AA;

Query Match 89.1%; Score 49; DB 9; Length 9;
Best Local Similarity 88.9%; Pred. No. 2.1e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 IETWFLRHP 9
Db 1 IETWILRHP 9

RESULT 4
ADW12594
ID ADW12594 standard; peptide; 21 AA.
XX AC ADW12594;
XX DT 24-MAR-2005 (first entry)
XX DE M20-40/DEN-2 mutant protein.
XX KW Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX KW DEN; dengue; mutant; mutein.
XX OS Dengue virus.
XX OS US2004266987-A1.
XX PN 30-DEC-2004.
XX PF 30-JUN-2003; 2003US-00608029.
XX PR 30-JUN-2003; 2003US-00608029.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX DR WPI; 2005-047647/05.
XX PT New isolated and purified Apoptom peptide comprises 9 amino acids, useful
XX as a vaccine for preventing or treating pathological conditions from non-
XX specific febrile illnesses to severe hemorrhagic manifestations or
XX encephalitic syndromes.
XX PS Example 3; Fig 4; 30pp; English.
XX CC The present invention relates to an isolated and purified Apoptom
XX peptide. The invention is useful as a vaccine for the prevention and
XX treatment of pathological conditions from non-specific febrile illnesses
XX to severe hemorrhagic manifestations, encephalitic syndromes and these
XX pathological conditions are linked to Flavivirus infection or cancers.
XX The invention is also useful in gene therapy. The present sequence is a
XX M32-40/DEN (dengue)-2 mutant protein.
XX SQ Sequence 9 AA;

Query Match 89.1%; Score 49; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 IETWFLRHP 9
Db 13 IETWILRHP 21

RESULT 5
ADW12593
ID ADW12593 standard; peptide; 32 AA.
XX AC ADW12593;
XX DT 24-MAR-2005 (first entry)
XX DE M10-40/DEN-2 mutant protein.
XX KW Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX KW DEN; dengue; mutant; mutein.
XX OS Dengue virus.
XX OS US2004266987-A1.
XX PN 30-DEC-2004.
XX PF 30-JUN-2003; 2003US-00608029.
XX PR 30-JUN-2003; 2003US-00608029.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX DR WPI; 2005-047647/05.
XX PT New isolated and purified Apoptom peptide comprises 9 amino acids, useful
XX as a vaccine for preventing or treating pathological conditions from non-
XX specific febrile illnesses to severe hemorrhagic manifestations or
XX encephalitic syndromes.
XX PS Example 3; Fig 4; 30pp; English.
XX CC The present invention relates to an isolated and purified Apoptom
XX peptide. The invention is useful as a vaccine for the prevention and
XX treatment of pathological conditions from non-specific febrile illnesses
XX to severe hemorrhagic manifestations, encephalitic syndromes and these
XX pathological conditions are linked to Flavivirus infection or cancers.
XX The invention is also useful in gene therapy. The present sequence is a
XX M10-40/DEN (dengue)-2 mutant protein.
XX SQ Sequence 32 AA;

```

```

Query Match      89.1%; Score 49; DB 9; Length 32;
Best Local Similarity 88.9%; Pred. No. 0.17;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
   ||| |||
Db 24 IETWILRHP 32

RESULT 6
ADW12576
ID ADW12576 standard; peptide; 39 AA.
AC ADW12576;
XX DT
XX 24-MAR-2005 (first entry)
XX DE M1-40/DEN-2 protein.
XX KW Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX KW DEN; dengue.
XX OS Dengue virus.
XX US2004266987-A1.
XX PN
XX 30-DEC-2004.
XX 30-JUN-2003; 2003US-00608029.
XX 30-JUN-2003; 2003US-00608029.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX WPI; 2005-047647/05.
XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
XX as a vaccine for preventing or treating pathological conditions from non-
XX specific febrile illnesses to severe hemorrhagic manifestations or
XX encephalitic syndromes.
XX Example 3; SEQ ID NO 23; 30pp; English.
XX The present invention relates to an isolated and purified ApoptoM
XX peptide. The invention is useful as a vaccine for the prevention and
XX treatment of pathological conditions from non-specific febrile illnesses
XX to severe hemorrhagic manifestations, encephalitic syndromes and these
XX pathological conditions are linked to Flavivirus infection or cancers.
XX The invention is also useful in gene therapy. The present sequence is a
XX M1-40/DEN (dengue)-2 protein.
XX Sequence 39 AA;

Query Match      89.1%; Score 49; DB 9; Length 39;
Best Local Similarity 88.9%; Pred. No. 0.21;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
   ||| |||
Db 31 IETWILRHP 39

RESULT 7
AAE17432
ID AAE17432 standard; peptide; 40 AA.
XX AC AAE17432;
XX 29-AUG-2003 (revised)
XX 18-APR-2002 (first entry)

```

---

```

XX Dengue (DEN)-2 virus M ectodomain.
DE Dengue virus; prM glycoprotein; E glycoprotein; apoptosis; virucide;
KW cancer; flavivirus infection; cytostatic; DEN-2 M ectodomain.
XX Dengue virus; 2.
OS WO200196376-A2.
XX 20-DEC-2001.
XX 18-JUN-2001; 2001WO-IB001570.
XX 16-JUN-2000; 2000US-0212129P.
XX (INSP ) INST PASTEUR.
XX Despres P, Courageot M, Deubel V, Catteau A;
XX WPI; 2002-139706/18.
XX Novel apoptosis inducing polypeptide fragments of Dengue virus-1 or 2 M
XX protein, useful for inducing apoptosis in a cell of a human patient
XX suffering from cancer or flavivirus infection.
XX Claim 9; Fig 12; 45pp; English.
XX The invention relates to pro-apoptotic fragments of the Dengue virus
XX (DEN) prM and E glycoproteins, methods for screening molecules capable of
XX inducing apoptosis and methods of inducing apoptosis in a cell. The
XX invention particularly relates to DEN-1 M (a membrane protein anchored in
XX envelope surrounding the nucleocapsid of the virus) ectodomain sequence,
XX Den-1-C amino acid sequence and DEN-2 M ectodomain sequence. Sequences of
XX the invention are useful for inducing apoptosis in a cell of a patient
XX suffering from cancer or flavivirus infection. They are also useful for
XX screening molecules which inhibit apoptosis. The present sequence is DEN-
XX 2 virus M ectodomain. (Updated on 29-AUG-2003 to standardise OS field)
XX Sequence 40 AA;

Query Match      89.1%; Score 49; DB 5; Length 40;
Best Local Similarity 88.9%; Pred. No. 0.22;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
   ||| |||
Db 32 IETWILRHP 40

RESULT 8
ADW12578
ID ADW12578 standard; peptide; 40 AA.
XX AC ADW12578;
XX 24-MAR-2005 (first entry)
XX DE M1-40/YF.17D (T34, I36, I37, H39) mutant protein.
XX KW Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX KW YF; yellow fever; mutant; mutein.
XX OS Yellow fever virus.
XX US2004266987-A1.
XX PN 30-DEC-2004.
XX 30-JUN-2003; 2003US-00608029.
XX 30-JUN-2003; 2003US-00608029.

```



XX (INSP ) INST PASTEUR.  
 XX Despres P, Catteau A;  
 XX WPI; 2005-047647/05.  
 XX New isolated and purified Apoptom peptide comprises 9 amino acids, useful  
 PT as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.  
 XX Example 3; SEQ ID NO 25; 30pp; English.  
 XX The present invention relates to an isolated and purified Apoptom  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC M1-40/YF (yellow fever).17D (T34, I36, I37, H39) mutant protein.  
 XX  
 SQ Sequence 40 AA;  
 Query Match 89.1%; Score 49; DB 9; Length 40;  
 Best Local Similarity 88.9%; Pred. No. 0.22;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 IETWFLRHP 9  
 |||||  
 DB 32 IETWILRHP 40  
 RESULT 9  
 AAEL17433  
 ID AAEL17433 standard; protein; 48 AA.  
 XX  
 AC AAEL17433;  
 DT 18-APR-2002 (first entry)  
 XX  
 DE (95-114)EGFP(206-245)DEN-2 fusion protein.  
 XX  
 KW Dengue virus; pRM glycoprotein; E glycoprotein; apoptosis; virucide;  
 KW cancer; flavivirus infection; cytostatic; EGFP; DEN-2 protein;  
 KW enhanced green fluorescent protein; fusion protein; M ectodomain.  
 XX  
 OS Dengue virus; 2.  
 OS Dengue virus; 1.  
 OS Unidentified.  
 OS Chimeric.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 13..44  
 FT /note= "Encoded by GTATC"  
 XX  
 FN WO200196376-A2.  
 XX  
 PD 20-DEC-2001.  
 XX  
 PF 18-JUN-2001; 2001WO-IB001570.  
 XX  
 PR 16-JUN-2000; 2000US-0212129P.  
 XX  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Despres P, Courageot M, Deubel V, Catteau A;  
 XX  
 DR WPI; 2002-139706/18.  
 DR N-PSDB; AAD27335.  
 XX  
 XX Novel apoptosis inducing polypeptide fragments of Dengue virus-1 or 2 M  
 PT protein, useful for inducing apoptosis in a cell of a human patient

PT suffering from cancer or flavivirus infection.  
 XX  
 PS Claim 42; Fig 11; 45pp; English.  
 XX  
 CC The invention relates to pro-apoptotic fragments of the Dengue virus  
 CC (DEN) pRM and E glycoproteins, methods for screening molecules capable of  
 CC inducing apoptosis and methods of inducing apoptosis in a cell. The  
 CC invention particularly relates to DEN-1 M (a membrane protein anchored in  
 CC envelope surrounding the nucleocapsid of the virus) ectodomain sequences,  
 CC Den-1-C amino acid sequence and DEN-2 M ectodomain sequence. Sequences of  
 CC the invention are useful for inducing apoptosis in a cell of a patient  
 CC suffering from cancer or flavivirus infection. They are also useful for  
 CC screening molecules which inhibit apoptosis. The present sequence is (95-  
 CC 114)EGFP(206-245)DEN-2 fusion protein construct. This construct comprises  
 CC 95-114 of the C-terminus of the C-protein of the DEN-1 virus strain BR/90  
 CC fused to the N-terminus of the enhanced green fluorescent protein (EGFP) and  
 CC DEN-2 virus strain Jamaica M ectodomain (DEN-2 polypeptide) fused to the  
 CC C-terminus of the EGFP sequence  
 XX  
 SQ Sequence 48 AA;  
 Query Match 89.1%; Score 49; DB 5; Length 48;  
 Best Local Similarity 88.9%; Pred. No. 0.27;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 IETWFLRHP 9  
 |||||  
 DB 40 IETWILRHP 48  
 RESULT 10  
 ADN37497  
 ID ADN37497 standard; protein; 167 AA.  
 XX  
 AC ADN37497;  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Dengue virus C15/truncated pRM antigen fusion protein - SEQ ID 122.  
 XX  
 KW virucide; Flavivirus; arboviruses group B; gene therapy; truncated pRM;  
 KW capsid.  
 XX  
 OS Dengue virus.  
 OS  
 XX WO2003102166-A2.  
 PD 11-DEC-2003.  
 XX  
 PF 26-FEB-2003; 2003WO-US005918.  
 XX  
 PR 26-FEB-2002; 2002US-0360030P.  
 XX  
 PA (MAXY-) MAXYGEN INC.  
 XX  
 PI Apt D, Punnonen J, Brinkman AM;  
 XX  
 DR WPI; 2004-043106/04.  
 XX  
 PT New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX  
 PS Disclosure; SEQ ID NO 122; 409pp; English.  
 XX  
 CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current

CC sequence is that of a Dengue virus C15/truncated prM antigen fusion  
 CC protein of the invention which comprises the C-terminal 15 amino acids of  
 CC the capsid protein fused to a truncated form of the prM protein lacking  
 CC the C-terminal 15 amino acids.

XX SQ Sequence 167 AA;

Query Match 89.1%; Score 49; DB 8; Length 167;  
 Best Local Similarity 88.9%; Pred. No. 1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9  
 |||||  
 Db 139 IETWILRHP 147

#### RESULT 11

ADN37493  
 ID ADN37493 standard; protein; 171 AA.

XX AC ADN37493;

XX DT 17-JUN-2004 (first entry)

XX DE Dengue virus type 2 (DEN-2) C15/truncated prM antigen fusion protein.

XX KW virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;  
 XX capsid; DEN-2.

XX OS Dengue virus type 2.

XX PN WO2003102166-A2.

XX PD 11-DEC-2003.

XX PF 26-FEB-2003; 2003WO-US005918.

XX PR 26-FEB-2002; 2002US-0360030P.

XX PA (MAXY-) MAXYGEN INC.

XX PI Apt D, Punnonen J, Brinkman AM;

XX DR WPI; 2004-043106/04.

XX PT New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

XX PS Disclosure; SEQ ID NO 118; 409pp; English.

XX CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of a Dengue virus type 2 (DEN-2) C15/truncated prM  
 CC antigen fusion protein of the invention which comprises the C-terminal 15  
 CC amino acids of the capsid protein fused to a truncated form of the prM  
 CC protein lacking the C-terminal 15 amino acids.

XX SQ Sequence 171 AA;

Query Match 89.1%; Score 49; DB 8; Length 171;  
 Best Local Similarity 88.9%; Pred. No. 1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9  
 |||||  
 Db 143 IETWILRHP 151

#### RESULT 12

ADN37496

ID ADN37496 standard; protein; 171 AA.

XX AC ADN37496;

XX DT 17-JUN-2004 (first entry)

XX DE Dengue virus C15/truncated prM antigen fusion protein - SEQ ID 121.

XX KW virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;  
 XX capsid.

XX OS Dengue virus.

XX PN WO2003102166-A2.

XX PD 11-DEC-2003.

XX PF 26-FEB-2003; 2003WO-US005918.

XX PR 26-FEB-2002; 2002US-0360030P.

XX PA (MAXY-) MAXYGEN INC.

XX PI Apt D, Punnonen J, Brinkman AM;

XX DR WPI; 2004-043106/04.

XX PT New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

XX PS Disclosure; SEQ ID NO 121; 409pp; English.

XX CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of a Dengue virus C15/truncated prM antigen fusion  
 CC protein of the invention which comprises the C-terminal 15 amino acids of  
 CC the capsid protein fused to a truncated form of the prM protein lacking  
 CC the C-terminal 15 amino acids.

XX SQ Sequence 171 AA;

Query Match 89.1%; Score 49; DB 8; Length 171;  
 Best Local Similarity 88.9%; Pred. No. 1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9  
 |||||  
 Db 143 IETWILRHP 151

#### RESULT 13

AAW75410

ID AAW75410 standard; peptide; 635 AA.

XX AC AAW75410;

XX DT 17-OCT-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 02-MAR-1999 (first entry)

XX DE Fusion protein PD30 contains Dengue virus epitope.

Dengue virus; fusion protein; P64K; Neisseria meningitidis; epitope; antibody; diagnosis; Flavivirus; infection; vaccine.

Dengue virus.  
Neisseria meningitidis.  
Chimeric.

WO9831814-A1.

23-JUL-1998.

13-JAN-1998; 98WO-CU000001.

15-JAN-1997; 97CU-00000013.

(CIGB-) CIGB CENT ING GENETICA & BIOTECNOLOGIA.  
(IPKM-) IPK INST MEDICINA TROPICAL KOURI PEDRO.

Vazquez Ramudo S, Guzman Tirado G, Guillen Nieto GE, Pardo Iazo OL; Chinae Santiago G, Perez Diaz AB, Pupo Antunez M, Rodriguez Roche R; Reyes Acosta O, Garay Perez HE, Padron Palomares G, Alvarez Vera M; Morier Diaz L, Perez Insueta O, Pelegrino Martinez De La Coterri Pedro; WPI; 1998-414111/35.

New peptide(s) and fusion proteins useful for diagnosis and treatment of flavivirus infection - contain cross-reactive epitopes from Dengue virus pre-M/M protein and can induce neutralising antibodies.

Claim 7; Page 28-29; 64pp; Spanish.

This protein represents a fusion protein comprising an M protein epitope from Dengue virus type 2 inserted into the P64K protein from Neisseria meningitidis. Synthetic peptides based on the Dengue virus epitope sequences (AAW75404-W75408) and fusion proteins can be used to raise antibodies. The peptides, protein and antibodies are all useful for diagnosis and treatment of Flavivirus infection, e.g. in vaccines.  
(Updated on 25-MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise OS field)

Sequence 635 AA;

Query Match 89.1%; Score 49; DB 2; Length 635;  
Best Local Similarity 88.9%; Pred. No. 4.3;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9  
|||||  
Db 74 IETWILRHP 82

RESULT 14

ADN37628

ID ADN37628 standard; protein; 675 AA.

AC ADN37628;

17-JUN-2004 (first entry)

Dengue virus C15/prM/E part codon-optimised antigen fusion protein 2.

virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; human codon-optimised; prM; envelope; capsid.

Dengue virus.  
Synthetic.

WO2003102166-A2.

11-DEC-2003.

26-FEB-2003; 2003WO-US005918.

26-FEB-2002; 2002US-0360030P.  
(MAXY-) MAXYGEN INC.

Apt D, Punnonen J, Brinkman AM;  
WPI; 2004-043106/04.  
N-PSDB; ADN37632.

New recombinant or synthetic polypeptides and polynucleotides useful for diagnosing, preventing or treating diseases associated with flaviviruses, including dengue viruses.

Example 28; SEQ ID NO 253; 409pp; English.

The invention relates to a novel recombinant or synthetic polypeptide comprising an amino acid sequence that has at least about 90% sequence identity to any of the 20 fully defined amino acid sequences given in the specification. The polypeptide of the invention demonstrates virucide activity and may be useful for inducing an immune response to Flaviviruses (arboviruses group B), including Dengue viruses, as well as in detecting and/or diagnosing the presence of antibodies against the Dengue virus serotypes in a sample and for gene therapy. The current sequence is that of the Dengue virus C15/prM/E partially human codon-optimised antigen fusion protein of the invention which comprises 15 amino acids of the capsid (C) protein fused to the full-length partially codon-optimised prM protein and envelope (E) protein.

Sequence 675 AA;

Query Match 89.1%; Score 49; DB 8; Length 675;  
Best Local Similarity 88.9%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9  
|||||  
Db 139 IETWILRHP 147

RESULT 15

ADN37518

ID ADN37518 standard; protein; 675 AA.

AC ADN37518;

17-JUN-2004 (first entry)

Dengue virus C15/prM/E antigen fusion protein - SEQ ID 143.

virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM; envelope; capsid.

Dengue virus.

WO2003102166-A2.

11-DEC-2003.

26-FEB-2003; 2003WO-US005918.

26-FEB-2002; 2002US-0360030P.

(MAXY-) MAXYGEN INC.

Apt D, Punnonen J, Brinkman AM;  
WPI; 2004-043106/04.

New recombinant or synthetic polypeptides and polynucleotides useful for diagnosing, preventing or treating diseases associated with flaviviruses, including dengue viruses.

Claim 40; SEQ ID NO 143; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.  
 XX SQ Sequence 675 AA;

Query Match 89.1%; Score 49; DB 8; Length 675;  
 Best Local Similarity 88.9%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9  
 |||||  
 Db 139 IETWILRHP 147

RESULT 16  
 ADN37612  
 ID ADN37612 standard; protein; 675 AA.  
 AC ADN37612;  
 XX 17-JUN-2004 (first entry)  
 DT Dengue virus C15/prM/E antigen fusion protein - SEQ ID 237.  
 DE virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;  
 KW envelope; capsid.  
 KW Dengue virus.  
 OS WO2003102166-A2.  
 PN 11-DEC-2003.  
 PD 26-FEB-2003; 2003WO-US005918.  
 PF 26-FEB-2002; 2002US-0360030P.  
 PR (MAXY-) MAXYGEN INC.  
 PA Apt D, Punnonen J, Brinkman AM;  
 PI WPI; 2004-043106/04.  
 DR New recombinant or synthetic polypeptides and polynucleotides useful for  
 XX diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX Claim 40; SEQ ID NO 237; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.  
 XX SQ Sequence 675 AA;

Query Match 89.1%; Score 49; DB 8; Length 675;  
 Best Local Similarity 88.9%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9  
 |||||  
 Db 139 IETWILRHP 147

RESULT 17  
 ADN37626  
 ID ADN37626 standard; protein; 675 AA.  
 AC ADN37626;  
 XX 17-JUN-2004 (first entry)  
 DT Dengue virus C15/prM/E part codon-optimised antigen fusion protein 1.  
 DE virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E;  
 KW human codon-optimised; prM; envelope; capsid.  
 XX Dengue virus.  
 OS Synthetic.  
 OS WO2003102166-A2.  
 PN 11-DEC-2003.  
 PD 26-FEB-2003; 2003WO-US005918.  
 PF 26-FEB-2002; 2002US-0360030P.  
 PR (MAXY-) MAXYGEN INC.  
 PA Apt D, Punnonen J, Brinkman AM;  
 PI WPI; 2004-043106/04.  
 DR N-PSDB; ADN37630.  
 XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX Claim 40; SEQ ID NO 251; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E partially human codon-  
 CC optimised antigen fusion protein of the invention which comprises 15  
 CC amino acids of the capsid (C) protein fused to the full-length partially  
 CC codon-optimised prM protein and envelope (E) protein.  
 XX SQ Sequence 675 AA;

Query Match 89.1%; Score 49; DB 8; Length 675;  
 Best Local Similarity 88.9%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9  
 |||||  
 Db 139 IETWILRHP 147

RESULT 18  
 AAW75411  
 ID AAW75411 standard; peptide; 677 AA.

```

XX AC AAW75411;
XX 17-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 02-MAR-1999 (first entry)
XX DE Fusion protein PD34 contains Dengue virus epitope.
XX KW Dengue virus; fusion protein; P64K; Neisseria meningitidis; epitope;
XX KW antibody; diagnosis; Flavivirus; infection; vaccine.
XX OS Dengue virus.
XX OS Neisseria meningitidis.
XX OS Chimeric.
XX FN WO9831814-A1.
XX XX 23-JUL-1998.
XX PF 13-JAN-1998; 98WO-CU000001.
XX XX 15-JAN-1997; 97CU-00000013.
XX PA (CIGB-) CIGB CENT ING GENETICA & BIOTECNOLOGIA.
XX PA (IPKM-) IPK INST MEDICINA TROPICAL KOURI PEDRO.
XX PI Vazquez Ramudo S, Guzman Tirado G, Guillen Nieto GE, Pardo Lazo OL;
XX PI China Santiago G, Perez Diaz AB, Pupo Antunez M, Rodriguez Roche R;
XX PI Reyes Acosta O, Garay Perez HE, Padron Palomares G, Alvarez Vera M;
XX PI Morier Diaz L, Perez Insuiza O, Pelegrino Martinez De La Coterri Pedro;
XX DR WPI; 1998-414111/35.
XX XX
XX XX New peptide(s) and fusion proteins useful for diagnosis and treatment of
XX PT flavivirus infection - contain cross-reactive epitopes from Dengue virus
XX PT pre-W/M protein and can induce neutralising antibodies.
XX PS Claim 7; Page 30-32; 64pp; Spanish.
XX CC
XX CC This protein represents a fusion protein comprising an M protein epitope
XX CC from Dengue virus type 4 inserted into the P64K protein from Neisseria
XX CC meningitidis. Synthetic peptides based on the Dengue virus epitope
XX CC sequences (AAW75404-W75408) and fusion proteins can be used to raise
XX CC antibodies. The peptides, protein and antibodies are all useful for
XX CC diagnosis and treatment of Flavivirus infection, e.g. in vaccines.
XX CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 677 AA;
XX Query Match 89.1%; Score 49; DB 2; Length 677;
XX Best Local Similarity 88.9%; Pred. No. 4.6;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 IETWFLRHP 9
XX DB 116 IETWILRHP 124
XX RESULT 19
XX ADN37613
XX ID ADN37613 standard; protein; 677 AA.
XX AC ADN37613;
XX XX
XX DT 17-JUN-2004 (first entry)
XX DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 238.
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;
XX KW envelope; capsid.
XX XX
XX OS Dengue virus.
XX OS WO2003102166-A2.
XX XX 11-DEC-2003.
XX PF 26-FEB-2003; 2003WO-US005918.
XX PR 26-FEB-2002; 2002US-0360030P.
XX PA (MAXY-) MAXYGEN INC.
XX PI Apt D, Punnonen J, Brinkman AM;
XX DR WPI; 2004-043106/04.
XX XX
XX XX The invention relates to a novel recombinant or synthetic polypeptide
XX CC comprising an amino acid sequence that has at least about 90% sequence
XX CC identity to any of the 20 fully defined amino acid sequences given in the
XX CC specification. The polypeptide of the invention demonstrates virucide
XX CC activity and may be useful for inducing an immune response to
XX CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
XX CC in detecting and/or diagnosing the presence of antibodies against the
XX CC Dengue virus serotypes in a sample and for gene therapy. The current
XX CC invention is that of the Dengue virus C15/prM/E antigen fusion protein of
XX CC the invention which comprises 15 amino acids of the capsid (C) protein
XX CC fused to the full-length prM protein and envelope (E) protein.
XX SQ Sequence 677 AA;
XX Query Match 89.1%; Score 49; DB 8; Length 677;
XX Best Local Similarity 88.9%; Pred. No. 4.6;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 IETWFLRHP 9
XX DB 139 IETWILRHP 147
XX RESULT 20
XX ADN37603
XX ID ADN37603 standard; protein; 681 AA.
XX XX
XX AC ADN37603;
XX XX
XX DT 17-JUN-2004 (first entry)
XX DE Dengue virus type 2 Den-2C15/prM/E antigen fusion protein.
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; DEN-2;
XX KW Den-2C15/prM/E; prM; envelope; capsid.
XX XX
XX OS Dengue virus type 2.
XX OS WO2003102166-A2.
XX XX 11-DEC-2003.
XX PF 26-FEB-2003; 2003WO-US005918.
XX PR 26-FEB-2002; 2002US-0360030P.
XX PA (MAXY-) MAXYGEN INC.
XX PI Apt D, Punnonen J, Brinkman AM;
XX DR WPI; 2004-043106/04.

```

CC Dengue virus serotypes in a sample and for gene therapy. The current  
CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
CC the invention which comprises 15 amino acids of the capsid (C) protein  
CC fused to the full-length prM protein and envelope (E) protein.  
XX  
XX  
SQ Sequence 681 AA;  
  
Query Match 89.1%; Score 49; DB 8; Length 681;  
Best Local Similarity 88.9%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 IETWFLRHP 9  
Db 143 IETWILRHP 151  
|||||  
  
RESULT 22  
ABP57874  
ID ABP57874 standard; protein; 685 AA.  
XX  
AC ABP57874;  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Plasmid pCBD2-14-6 containing dengue-2 virus prM and E.  
XX  
KW Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;  
KW pCBD2-14-6; dengue virus; DEN-2.  
XX  
OS Unidentified.  
OS Dengue-2 virus.  
OS Chimeric.  
XX  
XX WO200281754-A1.  
PN  
XX  
PD 17-OCT-2002.  
XX  
PF 04-APR-2002; 2002WO-US010764.  
XX  
XX 04-APR-2001; 2001US-00826115.  
PR  
XX (USSH ) US DBPT HEALTH & HUMAN SERVICES.  
PA  
XX Chang GJ;  
PI  
XX WPI; 2003-058572/05.  
DR N-PSDB; ABV77547.  
XX  
XX Novel isolated nucleic acid useful as vaccine for preventing flavivirus  
XX infection, comprises transcriptional unit encoding signal sequence of one  
XX flavivirus and immunogenic flavivirus antigen of a second flavivirus.  
PS Example 20; Page 157-158; 174pp; English.  
XX  
XX The invention relates to a novel nucleic acid comprising a  
XX transcriptional unit encoding a signal sequence of a structural protein  
XX of a first flavivirus and an immunogenic flavivirus antigen of a second  
XX flavivirus, where the transcriptional unit directs the synthesis of the  
XX antigen. The polynucleotide of the invention has virucide activity, and  
XX acts as a vaccine. A composition of the invention is useful for  
XX immunising a subject against infection by a flavivirus. The  
XX polynucleotide is useful as a vaccine for preventing flavivirus  
XX infection. The sequence represents plasmid pCBD2-14-6, which contains  
XX dengue-2 virus (DEN-2) prM and E proteins  
XX  
SQ Sequence 685 AA;  
  
Query Match 89.1%; Score 49; DB 6; Length 685;  
Best Local Similarity 88.9%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 IETWFLRHP 9  
|||||  
  
CC New recombinant or synthetic polypeptides and polynucleotides useful for  
CC diagnosing, preventing or treating diseases associated with flaviviruses,  
CC including dengue viruses.  
XX  
XX Claim 38; SEQ ID NO 228; 409pp; English.  
PS  
XX The invention relates to a novel recombinant or synthetic polypeptide  
CC comprising an amino acid sequence that has at least about 90% sequence  
CC identity to any of the 20 fully defined amino acid sequences given in the  
CC specification. The polypeptide of the invention demonstrates virucide  
CC activity and may be useful for inducing an immune response to  
CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
CC in detecting and/or diagnosing the presence of antibodies against the  
CC Dengue virus serotypes in a sample and for gene therapy. The current  
CC sequence is that of the Dengue virus type 2 (DEN-2) Den-2C15/prM/E  
CC antigen fusion protein of the invention which comprises 15 amino acids of  
CC the capsid (C) protein fused to the full-length prM protein and envelope  
CC (E) protein.  
XX  
XX  
SQ Sequence 681 AA;  
  
Query Match 89.1%; Score 49; DB 8; Length 681;  
Best Local Similarity 88.9%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 IETWFLRHP 9  
Db 143 IETWILRHP 151  
|||||  
  
RESULT 21  
ADN37517  
ID ADN37517 standard; protein; 681 AA.  
XX  
AC ADN37517;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 142.  
XX  
XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;  
KW envelope; capsid.  
XX  
OS Dengue virus.  
XX  
XX WO2003102166-A2.  
PN  
XX  
PD 11-DEC-2003.  
XX  
XX 26-FEB-2003; 2003WO-US005918.  
PF  
XX 26-FEB-2002; 2002US-0360030P.  
PR  
XX (MAXY-) MAXYGEN INC.  
PA  
XX Apt D, Punnnonen J, Brinkman AM;  
PI  
XX WPI; 2004-043106/04.  
DR  
XX New recombinant or synthetic polypeptides and polynucleotides useful for  
XX diagnosing, preventing or treating diseases associated with flaviviruses,  
XX including dengue viruses.  
PS Claim 40; SEQ ID NO 142; 409pp; English.  
XX  
XX The invention relates to a novel recombinant or synthetic polypeptide  
CC comprising an amino acid sequence that has at least about 90% sequence  
CC identity to any of the 20 fully defined amino acid sequences given in the  
CC specification. The polypeptide of the invention demonstrates virucide  
CC activity and may be useful for inducing an immune response to  
CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
CC in detecting and/or diagnosing the presence of antibodies against the

```

Db      147 IETWILRHP 155

RESULT 23
ABP57876
ID   ABP57876 standard; protein; 685 AA.
AC   ABP57876;
DT   07-FEB-2003 (first entry)
DE   Plasmid pCB8D2-2J-2-9-1 protein product.
XX
KW   Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;
KW   pCB8D2-2J-2-9-1; Japanese encephalitis virus; dengue-2 virus; DEN-2.
OS   Unidentified.
OS   Synthetic.
PN   WO200281754-A1.
PD   17-OCT-2002.
PF   04-APR-2002; 2002WO-US010764.
PR   04-APR-2001; 2001US-00826115.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Chang GJ;
XX   WPI; 2003-058572/05.
XX   N-PSDB; ABV77549.
XX   Novel isolated nucleic acid useful as vaccine for preventing flavivirus
XX   infection, comprises transcriptional unit encoding signal sequence of one
XX   flavivirus and immunogenic flavivirus antigen of a second flavivirus.
XX   Example 20; Page 162-164; 174pp; English.
XX
CC   The invention relates to a novel nucleic acid comprising a
CC   transcriptional unit encoding a signal sequence of a structural protein
CC   of a first flavivirus and an immunogenic flavivirus antigen of a second
CC   flavivirus, where the transcriptional unit directs the synthesis of the
CC   antigen. The polynucleotide of the invention has virucide activity, and
CC   acts as a vaccine. A composition of the invention is useful for
CC   immunising a subject against infection by a flavivirus. The
CC   polynucleotide is useful as a vaccine for preventing flavivirus
CC   infection. The sequence represents plasmid pCB8D2-2J-2-9-1, which
CC   contains dengue-2 virus (DEN-2) prM, M and E, and Japanese encephalitis
CC   virus E proteins
XX
SQ   Sequence 685 AA;

Query Match      89.1%; Score 49; DB 6; Length 685;
Best Local Similarity 88.9%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 IETWFLRHP 9
        ||||| ||||
Db      147 IETWILRHP 155

RESULT 25
AAW09409
ID   AAW09409 standard; protein; 1127 AA.
XX
AC   AAW09409;
XX
DT   17-OCT-2003 (revised)
DT   19-MAY-1997 (first entry)
XX
DE   Dengue virus serotype 2 PR159/S1 polypeptide.
XX
KW   DEN-2; flavivirus; envelope protein; immunisation; vaccine.
XX
OS   Dengue virus; serotype 2.
XX
FH   Key      Location/Qualifiers
FH   Region   1..114
FH   FT       /label= Capsid
FT   Region   115..205
FT   FT       /label= Pre-membrane
FT   Region   206..280
FT   FT       /label= Membrane
FT   Region   281..775

```

```

FT FT          /label= Envelope
FT FT          296..395
FT FT          /label= Domain-B
FT FT Misc-difference 588
FT FT          /notes="amino acid residue 588 (Val) is Ile in wild-type
FT FT          PR159"
FT FT          776..1127
FT FT          /label= NS1
XX XX WO9637221-A1.
XX XX
XX XX 28-NOV-1996.
XX XX
XX XX 24-MAY-1996; 96WO-US007627.
XX XX
XX XX 24-MAY-1995; 95US-00448734.
XX XX 07-JUN-1995; 95US-00488807.
XX XX 10-JUN-1995; 95US-00500469.
XX XX
XX XX (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.
XX XX
XX XX Ivy JM, Nakano E, Clements D;
XX XX
XX XX WPI; 1997-020938/02.
XX XX N-PSDB; AAT47666.
XX XX
XX XX Sub:unit vaccine against flavivirus infection - contg. recombinant
XX XX envelope protein in secretable form, used for immunising against
XX XX flavivirus infection.
XX XX
XX XX Example 1; Fig 3A-D; 121pp; English.
XX XX
XX XX A polypeptide (AAW09409) comprises the capsid, pre-membrane, envelope and
XX XX NS1 proteins of dengue virus serotype 2 (DEN-2) variant PR159/S1. A
XX XX conservative mutation in the envelope protein may be involved in the
XX XX attenuation of this small-plaque, temp.- sensitive variant. Portions of
XX XX the envelope protein, esp. domain B, can be expressed in eukaryotic hosts
XX XX (see also AAW09410 and AAW09427-28) transfected with vectors
XX XX incorporating DEN-2 S1 cDNA (see also AAT47666). These polypeptides can
XX XX be used in novel subunit vaccines against viral infection, to raise
XX XX antibodies useful for passive immunisation, and for diagnosis of
XX XX infection. (Updated on 17-OCT-2003 to standardise OS field)
XX XX
XX XX Sequence 1127 AA;
XX XX
XX XX Query Match 89.1%; Score 49; DB 2; Length 1127;
XX XX Best Local Similarity 88.9%; Pred. No. 7.9;
XX XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX XX
XX XX QY 1 IETWFLRHP 9
XX XX ||||| |||||
XX XX 237 IETWILRHP 245
XX XX
XX XX RESULT 26
XX XX AAY05522
XX XX ID AAY05522 standard; protein; 1127 AA.
XX XX
XX XX AC AAY05522;
XX XX
XX XX 17-OCT-2003 (revised)
XX XX 05-JUL-1999 (first entry)
XX XX
XX XX Dengue virus serotype 2 PR159/S1 viral capsid, pprM, E, NS1.
XX XX
XX XX Flavivirus; envelope protein; vaccine; infection; diagnosis.
XX XX
XX XX Dengue virus; serotype 2.
XX XX
XX XX Key Location/Qualifiers
XX XX FT Protein 1..114
XX XX FT Protein /label= Capsid
XX XX FT Protein 115..205

```

```

FT FT Protein /label= PreMembrane
FT FT 206..280
FT FT Protein /label= Membrane
FT FT 280..1127
FT FT Protein /label= Envelope
XX XX WO9906068-A2.
XX XX
XX XX 11-FEB-1999.
XX XX
XX XX 27-JUL-1998; 98WO-US015447.
XX XX
XX XX 31-JUL-1997; 97US-00904227.
XX XX
XX XX (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.
XX XX
XX XX Ivy JM, Peters ID, Collier BG, McDonnell M, Harada KE;
XX XX WPI; 1999-153454/13.
XX XX N-PSDB; AAX25114.
XX XX
XX XX Recombinant dimeric flaviviral envelope vaccine - comprising a dimeric
XX XX 80%E protein, useful for protecting against flavivirus, especially dengue
XX XX virus infections.
XX XX
XX XX Example 1; Fig 3A-D; 60pp; English.
XX XX
XX XX This sequence is composed of the capsid, prM, envelope (E) and NS1
XX XX proteins of serotype 2 dengue virus DEN-2 strain PR159/S1. A vaccine for
XX XX protecting against flavivirus infection comprises a dimeric 80% E protein
XX XX that has been secreted as a recombinant protein from a eukaryotic cell.
XX XX 80% E indicates a C-terminally truncated flavivirus E protein. The
XX XX dimeric truncated E is formed: (1) by directly linking 2 tandem copies of
XX XX 80% E via a flexible tether; (2) via the formation of a leucine zipper
XX XX domain through the homodimeric association of 2 leucine zipper helices
XX XX each fused to the C-terminus of an 80% E molecule; or (3) via the
XX XX formation of a non-covalently associated four-helix bundle domain formed
XX XX upon association of two helix-turn-helix moieties attached to the C-
XX XX terminus of an 80% E molecule. Dimeric truncated DEN-2 E proteins are
XX XX efficiently secreted by recombinant cells, are easier to purify than
XX XX intracellular proteins, and generate a high titer neutralising antibody
XX XX response. The method is generally applicable to flaviviruses, in
XX XX particular dengue viruses such as DEN-2, where 80% E comprises amino
XX XX acids 1-395 of DEN-2 E. The products can also be used for diagnosis of
XX XX infection. (Updated on 17-OCT-2003 to standardise OS field)
XX XX
XX XX Sequence 1127 AA;
XX XX
XX XX Query Match 89.1%; Score 49; DB 2; Length 1127;
XX XX Best Local Similarity 88.9%; Pred. No. 7.9;
XX XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX XX
XX XX QY 1 IETWFLRHP 9
XX XX ||||| |||||
XX XX 237 IETWILRHP 245
XX XX
XX XX RESULT 27
XX XX ADL98086
XX XX ID ADL98086 standard; protein; 1127 AA.
XX XX
XX XX AC ADL98086;
XX XX
XX XX 18-NOV-2004 (first entry)
XX XX
XX XX Dengue virus, DEN-2, capsid/membrane/envelope/NS1 proteins.
XX XX
XX XX Dengue virus; DEN-2; Envelope protein; 80% E; membrane protein;
XX XX capsid protein; NS1 protein; Dengue haemorrhagic fever; DHF;
XX XX Dengue shock syndrome; DSS; flavivirus; vaccine.
XX XX
XX XX Dengue virus type 2; strain PR159/S1.

```



```

PN  US2003175304-A1.
XX
XX
PD  18-SEP-2003.
XX
XX
PF  20-SEP-2002; 2002US-00247960.
XX
XX
PR  31-JUL-1997; 97US-00904227.
PR  18-AUG-1999; 99US-00376463.
XX
XX
PA  (PETE/) PETERS I D.
PA  (COLL/) COLLIER B G.
PA  (MCDON/) MCDONELL M.
PA  (IVYJ/) IVY J M.
PA  (HARA/) HARADA K.
XX
XX
PI  Peters ID, Collier BG, McDonnell M, Ivy JM, Harada K;
XX
XX
WPI; 2003-898503/82.
DR  N-PSDB; ADL98085.
XX
XX
Vaccine useful for protection against dengue virus infection, comprises a
PT  dimeric 80% envelope, which has been secreted as a recombinantly produced
PT  protein from Drosophila Schneider cells.
XX
XX
PS  Example 1; Fig 3; 31pp; English.
XX
XX
The invention relates to a vaccine for protection against Flavivirus
CC  infection comprising a dimeric 80% envelope (E), which has been secreted
CC  as a recombinantly produced protein from Drosophila Schneider cells and
CC  which represents the N-terminal 80% portion of the protein from residue 1
CC  -395. Also included are a method for protecting a subject against a
CC  Flavivirus, an immunogenic polypeptide comprising a dimeric 80% E, an
CC  immunogenic composition for protection against Flavivirus infection
CC  comprising the immunogenic polypeptide and a carrier, an immunodiagnostic
CC  for detecting Flavivirus comprising the immunogenic polypeptide, a vector
CC  host recombinant DNA expression system, a DNA sequence encoding the
CC  immunogenic polypeptide and an immunodiagnostic kit for detecting
CC  Flavivirus in a test subject. The dimeric 80% E products are envelope
CC  proteins of serotypes comprising DEN-1, DEN-2, DEN-3 or DEN-4. The
CC  Flavivirus is a dengue virus. The 80% E protein is produced as a dimer by
CC  incorporating 2 different kinds of leucine zipper peptides or
CC  incorporating a helix-turn-helix peptide, to encourage dimerisation. The
CC  vaccine is useful for protection against dengue virus infection (e.g.
CC  Dengue haemorrhagic fever, DHF, and Dengue shock syndrome, DSS). The
CC  present sequence is encoded by the partial genomic sequence of the DEN-2
CC  strain PR159/S1 virus, and represents the capsid, membrane, envelope and
CC  NS1 proteins.
XX
XX
SQ  Sequence 1127 AA;

```

```

Query Match      89.1%; Score 49; DB 7; Length 1127;
Best Local Similarity 88.9%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy  1 IETWFLRHP 9
    |||||
Db  237 IETWILRHP 245

```

```

RESULT 28
ADQ28716
ID  ADQ28716 standard; protein; 1127 AA.
XX
XX
AC  ADQ28716;
XX
XX
DT  26-AUG-2004 (first entry)
XX
XX
DE  Dengue virus viral capsid, prM, E and NS1 gene polyprotein.
XX
XX
KW  virucide; vaccine; Flavivirus; dimeric 80%; Drosophila Schneider cell;
KW  immunogenic composition; multivalent immunodiagnostic; dengue virus;
KW  viral capsid; prM gene; E gene; NS1 gene.
XX
XX

```

```

OS  Dengue virus.
XX
XX
PN  US6749857-B1.
XX
XX
PD  15-JUN-2004.
XX
XX
PF  18-AUG-1999; 99US-00376463.
XX
XX
PR  31-JUL-1997; 97US-00904227.
XX
XX
PA  (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.
XX
XX
PI  Peters ID, Collier BG, McDonnell M, Ivy JM, Harada K;
XX
XX
WPI; 2004-438725/41.
DR  N-PSDB; ADQ28715.
XX
XX
New vaccines for preventing or diagnosing infections caused by dengue
PT  virus comprises a therapeutic amount of a dimeric 80% protein secreted
PT  from Drosophila Schneider cells.
XX
XX
PS  Example 1; SEQ ID NO 3; 47pp; English.
XX
XX
The invention describes a vaccine that generates a protective,
CC  neutralising antibody response to a Flavivirus in a murine host. The
CC  vaccine comprises a therapeutic amount of a dimeric 80% E, the dimeric
CC  80% E having been secreted as a recombinantly produced protein from
CC  Drosophila Schneider cells, and where 80% E represents the N-terminal 80%
CC  portion of the protein from residues 1-395. Also described are: an
CC  immunogenic polypeptide comprising the dimeric 80% E cited above; an
CC  immunogenic composition that generates a protective, neutralising
CC  antibody response to a Flavivirus in a murine host, comprising the above
CC  immunogenic polypeptide and a physiological carrier; a multivalent
CC  immunodiagnostic for the detection of Flavivirus, comprising at least 2
CC  of the above immunogenic polypeptides of at least 2 flavivirus serotypes;
CC  and an immunodiagnostic kit for the detection of Flavivirus in a test
CC  subject, comprising the above immunogenic or multivalent immunodiagnostic
CC  polypeptide, a suitable support phase coated with dimeric 80% E, and
CC  labeled antibodies immunoreactive to antibodies from the test subject.
CC  The composition is useful for preventing or diagnosing infections caused
CC  by dengue virus. This is the amino acid sequence of the polyprotein
CC  encoded by Dengue virus gene viral capsid, prM, E and NS1 genes for
CC  Dengue virus strain PR159/S1 used as the source of DEN-2 genes for the
CC  invention.
XX
XX
SQ  Sequence 1127 AA;

```

```

Query Match      89.1%; Score 49; DB 8; Length 1127;
Best Local Similarity 88.9%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy  1 IETWFLRHP 9
    |||||
Db  237 IETWILRHP 245

```

```

RESULT 29
AAE35314
ID  AAE35314 standard; protein; 3388 AA.
XX
XX
AC  AAE35314;
XX
XX
DT  28-MAY-2003 (first entry)
XX
XX
DE  Dengue virus type 2 strain rDEN2/4delta30 protein.
XX
XX
KW  Attenuation; growth; vaccine; infection; Dengue virus type 4.
XX
XX
OS  Dengue virus.
XX
XX
PN  W0200295075-A1.
XX
XX
PD  28-NOV-2002.

```

XX	22-MAY-2002; 2002WO-US016308.	FT	Modified-site	982	
PF		FT	Protein	/label= N-glycosylated	
XX		FT		1128..1345	
PR	22-MAY-2001; 2001US-0293049P.	FT	Modified-site	/label= NS2A	
XX		FT		1134	
XX	(USSH ) US DEPT HEALTH & HUMAN SERVICES.	FT	Modified-site	/label= N-glycosylated	
PA	(BLAN/) BLANEY J E.	FT		1174	
XX		FT	Modified-site	/label= N-glycosylated	
XX	Whitehead SS, Murphy BR, Hanley KA;	FT		1329	
XX		FT	Modified-site	/label= N-glycosylated	
XX	WPI; 2003-120809/11.	FT	Protein	1346..1474	
DR	N-PSDB; AADS3912.	FT		/label= NS2B	
XX		FT	Modified-site	1369	
XX	New mutated flavivirus, useful for fine tuning the attenuation and growth	FT		/label= N-glycosylated	
PT	characteristics of dengue virus vaccines for the prevention and/or	FT	Protein	1475..2093	
PT	treatment of dengue virus infection.	FT		/label= NS3	
XX		FT	Protein	2094..2243	
PS	Disclosure; Page 133-134; 246pp; English.	FT		/label= NS4A	
XX		FT	Protein	2244..2492	
XX	The present invention relates to novel mutated flaviviruses comprising a	FT		/label= NS4B	
CC	phenotype in which the viral genome is modified by introduction of a	FT	Modified-site	2301	
CC	mutation, singly or in combination, taken from mutations from recombinant	FT		/label= N-glycosylated	
CC	virus bearing Vero adaptation mutations, putative Vero cell adaptation	FT	Modified-site	2305	
CC	mutations of dengue type 4 virus (DEN4) or mutations known to attenuate	FT		/label= N-glycosylated	
CC	dengue type 4 virus. The methods and compositions of the invention are	FT	Modified-site	2457	
CC	useful for fine tuning the attenuation and growth characteristics of	FT		/label= N-glycosylated	
CC	dengue virus vaccines for the prevention and/or treatment of dengue virus	FT	Modified-site	2485	
CC	infection. The present sequence is Dengue virus type 4 strain	FT		/label= N-glycosylated	
CC	rDEN2/delta30 protein	FT	Protein	2493..3391	
XX		FT		/label= NS5	
SQ	Sequence 3388 AA;	FT	Modified-site	2644	
		FT		/label= N-glycosylated	
	Query Match 89.1%; Score 49; DB 6; Length 3388;	FT	Modified-site	2665	
	Best Local Similarity 88.9%; Pred. No. 26;	FT		/label= N-glycosylated	
	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	FT	Modified-site	2704	
		FT		/label= N-glycosylated	
QY	1 IETWFLRHP 9	FT	Modified-site	2714	
		FT		/label= N-glycosylated	
Db	237 IETWILRHP 245	XX		FR2654113-A.	
		PN			
		XX		10-MAY-1991.	
RESULT 30		XX			
AA013166		XX		09-NOV-1989; 89FR-00914724.	
ID	AA013166 standard; protein; 3391 AA.	XX			
XX		XX		09-NOV-1989; 89FR-00014724.	
AC	AA013166;	PR			
XX		XX		(INSP ) INST PASTEUR.	
25-MAR-2003 (revised)		PA			
21-NOV-1991 (first entry)		XX		Vincent D;	
Proteins encoded by entire Dengue 2 virus genome.		PI			
dengue virus; detection; consensus sequence; Flavivirus; PCR.		XX			
Dengue virus.		XX		WPI; 1991-225002/31.	
Key	Location/Qualifiers	DR		N-PSDB; AAQ12787.	
Peptide	116..205	XX		Detection and identification of Flaviviridae in biological sample - by	
	/label= prM	PT		amplifying consensus sequence then hybridisation opt. followed by typing,	
Modified-site	183	FT		e.g. sequencing amplified prod.	
	/label= N-glycosylated	PS		Disclosure; Fig 3; 24pp; French.	
Protein	206..280	XX			
	/label= M	CC		The dengue 2 virus is an example of a member of the Flaviviridae which	
Protein	281..775	CC		can be identified using the probe pair of the invention. A species-	
	/label= E	CC		specific sequence can be amplified using the claimed oligonucleotides as	
Modified-site	347	CC		primers in a PCR reaction (see AAQ12788 and AAQ12789). Other viruses	
	/label= N-glycosylated	CC		which can be identified include Japanese encephalitis virus and yellow	
Modified-site	433	CC		fever virus. All the dengue 2 virus proteins are encoded from an	
	/label= N-glycosylated	CC		uninterrupted genomic sequence. (Updated on 25-MAR-2003 to correct PR	
Protein	776..1127	CC		field.)	
	/label= NS1	XX		Sequence 3391 AA;	
Modified-site	905	SQ			
	/label= N-glycosylated			Query Match 89.1%; Score 49; DB 2; Length 3391;	
				Best Local Similarity 88.9%; Pred. No. 26;	

---

Matches	8;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	IETWFLRHP	9						
Db	237	IETWILRHP	245						

Search completed: August 31, 2006, 11:50:37  
Job time : 110.25 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: August 31, 2006, 11:43:31 ; Search time 17.25 Seconds  
(without alignments)  
50.200 Million cell updates/sec

Title: DENGUE\_SEROTYPE2

Perfect score: 55

Sequence: 1 ietwflrhp 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 80.\*

1: piri.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	89.1	166	2 S40144	premembrane protei
2	49	89.1	555	2 JQ1404	genome polyprotein
3	49	89.1	775	2 A48644	polyprotein - deng
4	49	89.1	3388	1 GNWVDP	genome polyprotein
5	49	89.1	3391	1 GNWV16	genome polyprotein
6	49	89.1	3391	1 GNWV26	genome polyprotein
7	49	89.1	3391	1 GNWV3A	genome polyprotein
8	49	89.1	3391	2 JS0219	polyprotein - deng
9	46	83.6	166	2 S09223	membrane protein -
10	46	83.6	166	2 S09225	membrane protein -
11	46	83.6	555	2 JQ1405	genome polyprotein
12	46	83.6	775	2 A47311	polyprotein(C, E,
13	46	83.6	792	2 C32401	genome polyprotein
14	46	83.6	792	2 B32401	genome polyprotein
15	46	83.6	792	2 A32401	genome polyprotein
16	46	83.6	1127	1 GNWVDP	genome polyprotein
17	46	83.6	1226	1 GNWVDP	genome polyprotein
18	46	83.6	3390	1 GNWV3	genome polyprotein
19	46	83.6	3396	1 A42551	genome polyprotein
20	43	78.2	665	2 PS0043	genome polyprotein
21	42	76.4	166	2 S09224	membrane protein -
22	41	74.5	205	2 S86085	hypothetical prote
23	41	74.5	205	2 A98238	hypothetical prote
24	41	74.5	422	2 A83184	probable protein m
25	40	72.7	205	2 I78665	hypothetical 23.0K
26	39	70.9	343	2 H95879	probable sugar ABC
27	39	70.9	3411	1 GNWVY	genome polyprotein
28	39	70.9	3411	1 GNWVYP	genome polyprotein
29	38	69.1	399	2 T49934	carboxypeptidase-1

30	69.1	560	1	VGBE14	glycoprotein gpV -
31	69.1	1155	2	B96761	probable protein k
32	67.3	144	2	B40098	colorectal cancer
33	67.3	301	2	C95872	hypothetical prote
34	67.3	427	2	F72389	conserved hypothet
35	67.3	533	2	T35722	probable transport
36	67.3	773	2	A47666	structural polypro
37	67.3	879	2	B70014	antibiotic synthet
38	67.3	1244	2	S37034	DNA-directed DNA p
39	67.3	1447	2	A54100	tumor suppressor p
40	67.3	1525	1	GNWVS5	genome polyprotein
41	67.3	2413	2	S34670	splicing factor PR
42	67.3	3386	1	GNWVDF	genome polyprotein
43	65.5	120	2	A97655	hypothetical prote
44	65.5	120	2	AG2878	conserved hypothet
45	65.5	217	2	A83146	lipote-protein li
46	65.5	266	2	S02510	nifM protein - Kle
47	65.5	267	2	A38442	probable tumor sup
48	65.5	295	2	AG0923	LysR-family regula
49	65.5	336	2	JE0215	nitrite reductase
50	65.5	343	2	G84711	hypothetical prote
51	65.5	360	2	JG0170	nitrite reductase
52	65.5	436	2	H69588	acetylornithine de
53	65.5	481	2	B83062	deoxyribodipyrimid
54	65.5	490	2	I41293	EcoE type I restri
55	65.5	516	1	FWSYG3	glycinin G5 precu
56	65.5	575	2	A49667	interleukin-10 rec
57	65.5	615	2	T47395	hypothetical prote
58	65.5	630	2	T02524	probable RING zinc
59	65.5	640	2	B32935	hypothetical prote
60	65.5	739	2	A90141	ATP-dependent heli
61	65.5	805	2	G87268	DNA gyrase subunit
62	65.5	826	2	B96712	probable receptor
63	65.5	1008	2	T12532	hypothetical prote
64	65.5	1332	2	F69732	PBSX prophage ORF
65	65.5	2236	1	QZFF	rudimentary protei
66	63.6	194	1	S49184	phosphinothricin N
67	63.6	297	2	F98323	hypothetical oxido
68	63.6	297	2	AH2959	aryl-alcohol dehyd
69	63.6	305	2	G84140	aryl-alcohol dehyd
70	63.6	306	2	B97315	aldo/keto reductas
71	63.6	322	2	B84908	hypothetical prote
72	63.6	328	2	B83321	conserved hypothet
73	63.6	333	2	T02690	hypothetical prote
74	63.6	384	2	S74774	hypothetical prote
75	63.6	403	2	S42532	hypothetical prote
76	63.6	417	1	VGBE1B	glycoprotein D pre
77	63.6	417	2	S35784	glycoprotein GD -
78	63.6	430	2	T28870	hypothetical prote
79	63.6	470	2	B96637	EcoA sytem protei
80	63.6	489	2	A47200	hypothetical prote
81	63.6	493	2	F86133	hypothetical prote
82	63.6	493	2	C91292	hypothetical prote
83	63.6	608	2	G82137	pvcA protein VC194
84	63.6	1042	2	T48801	hypothetical prote
85	63.6	1310	2	T40135	oxysterol-binding
86	63.6	1467	2	T23950	hypothetical prote
87	63.6	1496	2	T43274	dynein heavy chain
88	61.8	108	2	AB0756	hypothetical prote
89	61.8	208	2	T33341	hypothetical prote
90	61.8	216	2	H72291	hypothetical prote
91	61.8	224	2	B87657	conserved hypothet
92	61.8	225	2	G72291	hypothetical prote
93	61.8	235	2	AF0656	conserved hypothet
94	61.8	240	2	S75162	hypothetical prote
95	61.8	246	2	H70223	conserved hypothet
96	61.8	270	2	E64924	hypothetical prote
97	61.8	270	2	D85774	hypothetical prote
98	61.8	270	2	H90925	hypothetical prote
99	61.8	279	2	G84555	hypothetical prote
100	61.8	292	2	T08862	hypothetical prote



A:Molecule type: genomic RNA  
A:Residues: 1-3391 <BLO>  
A:Cross-references: UNIPROT:P29990; UNIPARC:UPI0000131DF5; GB:M84727; GB:M85259; NID:g32  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; capsid protein; envelope protein; nonstructural protein;  
F:1-114/Product: capsid protein C #status predicted <CPC>  
F:115-280/Product: membrane-associated protein M precursor #status predicted <MPP>  
F:115-280/Product: nonterminal signal sequence #status predicted <SIG>  
F:206-280/Product: membrane-associated protein M #status predicted <MPM>  
F:268-284/Domain: transmembrane #status predicted <TM1>  
F:281-775/Product: envelope protein E #status predicted <EPE>  
F:727-743/Domain: transmembrane #status predicted <TM2>  
F:757-773/Domain: transmembrane #status predicted <TM3>  
F:776-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1128-1345/Product: nonstructural protein NS2a #status predicted <NS2a>  
F:1346-1474/Product: nonstructural protein NS2b #status predicted <NS2b>  
F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1668-1675/Region: nucleotide-binding motif A (P-loop)  
F:1755-1760/Region: nucleotide-binding motif B  
F:1759-1762/Region: DEAH motif  
F:2094-2243/Product: nonstructural protein NS4a #status predicted <NS4a>  
F:2244-2491/Product: nonstructural protein NS4b #status predicted <NS4b>  
F:2492-3391/Product: nonstructural protein NS5 #status predicted <NS5>  
F:183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.1%; Score 49; DB 1; Length 3391;  
Best Local Similarity 88.9%; Pred. No. 5.4;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9  
|||||  
DB 237 IETWILRHP 245

RESULT 6  
GNWU26  
Genome polyprotein - dengue virus type 2 (strain 16681-PDK53)  
N:Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru  
tural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: dengue virus type 2  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 31-Dec-2004  
C:Accession: B42451  
R:Block, J.; McWilliam, S.M.; Butler, H.C.; Gibbs, A.J.; Weiller, B.L.; Heme  
Virolgy 187, 573-590, 1992  
A:Title: Comparison of a dengue-2 virus and its candidate vaccine derivative: sequence  
A:Reference number: A42451; MUID:92188532; PMID:1312269  
A:Accession: B42451  
A:Molecule type: genomic RNA  
A:Residues: 1-3391 <BLO>  
A:Cross-references: UNIPROT:P29991; UNIPARC:UPI0000131DF6; GB:M85259  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F:1-114/Product: capsid protein C #status predicted <CPC>  
F:102-118/Domain: transmembrane #status predicted <TM1>  
F:115-280/Product: membrane-associated protein M precursor #status predicted <MPP>  
F:115-280/Product: nonterminal signal sequence #status predicted <SIG>  
F:206-280/Product: membrane-associated protein M #status predicted <MPM>  
F:268-284/Domain: transmembrane #status predicted <TM3>  
F:281-775/Product: envelope protein E #status predicted <EPE>  
F:727-743/Domain: transmembrane #status predicted <TM4>  
F:757-773/Domain: transmembrane #status predicted <TM5>  
F:776-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1128-1345/Product: nonstructural protein NS2a #status predicted <NS2a>  
F:1282-1288/Domain: transmembrane #status predicted <TM6>  
F:1294-1310/Domain: transmembrane #status predicted <TM8>  
F:1346-1474/Product: nonstructural protein NS2b #status predicted <NS2b>  
F:1351-1367/Domain: transmembrane #status predicted <TM9>  
F:1373-1389/Domain: transmembrane #status predicted <TM9>  
F:1448-1464/Domain: transmembrane #status predicted <TM9>  
F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1668-1675/Region: nucleotide-binding motif A (P-loop)

F:1755-1760/Region: nucleotide-binding motif B  
F:1759-1762/Region: DEAH motif  
F:2094-2243/Product: nonstructural protein NS4a #status predicted <NS4a>  
F:2148-2164/Domain: transmembrane #status predicted <TM2>  
F:2174-2190/Domain: transmembrane #status predicted <TM2>  
F:2197-2213/Domain: transmembrane #status predicted <TM2>  
F:2227-2243/Domain: transmembrane #status predicted <TM2>  
F:2244-2491/Product: nonstructural protein NS4b #status predicted <NS4b>  
F:2352-2368/Domain: transmembrane #status predicted <TM2>  
F:2411-2427/Domain: transmembrane #status predicted <TM2>  
F:2492-3391/Product: nonstructural protein NS5 #status predicted <NS5>  
F:183,347,433,905,982,1134,1174,1329,2301,2305,2346,2387,2457,2485,2644,2665,2704,2714/B  
Query Match 89.1%; Score 49; DB 1; Length 3391;  
Best Local Similarity 88.9%; Pred. No. 5.4;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9  
|||||  
DB 237 IETWILRHP 245

RESULT 7  
GNWU2A  
Genome polyprotein - dengue virus type 2 (strain Jamaica)  
N:Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru  
tural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: dengue virus type 2  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 31-Dec-2004  
C:Accession: A94346; A94378; A25613; A29199  
R:Deubel, V.; Kinney, R.M.; Trent, D.W.  
Virolgy 155, 365-377, 1986  
A:Title: Nucleotide sequence and deduced amino acid sequence of the structural proteins  
A:Reference number: A94346; MUID:87071658; PMID:3024394  
A:Accession: A94346  
A:Molecule type: genomic RNA  
A:Residues: 1-791 <DB1>  
A:Cross-references: UNIPROT:P07564; UNIPARC:UPI00001710BB; GB:M15975  
R:Deubel, V.; Kinney, R.M.; Trent, D.W.  
Virolgy 165, 234-244, 1988  
A:Title: Nucleotide sequence and deduced amino acid sequence of the nonstructural protei  
A:Reference number: A94378; MUID:88265864; PMID:3388770  
A:Accession: A94378  
A:Molecule type: genomic RNA  
A:Residues: 792-3391 <DE2>  
A:Cross-references: UNIPARC:UPI0000174A05; GB:M20558  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F:2-114/Product: capsid protein C #status predicted <CPC>  
F:43-59/Domain: transmembrane #status predicted <TM1>  
F:101-117/Domain: transmembrane #status predicted <TM2>  
F:115-280/Product: membrane-associated protein M precursor #status predicted <MPP>  
F:115-280/Domain: nonterminal signal sequence #status predicted <SIG>  
F:206-280/Product: membrane-associated protein M #status predicted <MPM>  
F:268-284/Domain: transmembrane #status predicted <TM3>  
F:281-775/Product: envelope protein E #status predicted <EPE>  
F:727-743/Domain: transmembrane #status predicted <TM4>  
F:757-773/Domain: transmembrane #status predicted <TM5>  
F:776-1127/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1128-1345/Product: nonstructural protein NS2a #status predicted <NS2a>  
F:1346-1474/Product: nonstructural protein NS2b #status predicted <NS2b>  
F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1668-1675/Region: nucleotide-binding motif A (P-loop)  
F:1755-1760/Region: nucleotide-binding motif B  
F:1759-1762/Region: DEAH motif  
F:2094-2243/Product: nonstructural protein NS4a #status predicted <NS4a>  
F:2244-2491/Product: nonstructural protein NS4b #status predicted <NS4b>  
F:2492-3391/Product: nonstructural protein NS5 #status predicted <NS5>  
F:183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.1%; Score 49; DB 1; Length 3391;  
Best Local Similarity 88.9%; Pred. No. 5.4;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;





Quarry Match	83 68.	Score 15.	DB 3.	Length 703.
--------------	--------	-----------	-------	-------------

83.6%; Score 46; DB 2; Length 792;

Best Local Similarity 77.8%; Pred. No. 4; Matches 7; Conservative 1; Mismatches 0; Gaps 0;	
QY 1 IETWFLRHP 9 :	Db 237 VETWALRHP 245
RESULT 16 GNWVD2	
genome polyprotein - dengue virus type 2 (strain D2-04) (fragment) N;Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru C;Species: dengue virus type 2 C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 31-Dec-2004 C;Accession: JCI1007; JCI1005 R;Yang, P.Y.; Lam, S.K. Chinese J. Microbiol. Immunol. 11, 341-344, 1991 A;Title: The nucleotide and encoded amino acid sequences of the structural protein gene A;Reference number: JCI1007 A;Accession: JCI1007 A;Molecule type: genomic RNA A;Residues: 1-775 <YAN> A;Cross-references: UNIPROT:P10026; UNIPARC:UPI0000174A06 A;Note: the authors translated the codons TTA for residue 53 as Phe, AGT for residue 136 S as Arg, GGC for residue 266 as Ala, and CAG for residue 272 as Leu R;Yan, P.Y.; Kautner, I.M.; Koh, C.L.; Lam, S.K. Chinese J. Microbiol. Immunol. 11, 9-12, 1991 A;Title: Nucleotide and encoded amino acid sequences of the nonstructural protein NS1 ge A;Reference number: JCI1005 A;Accession: JCI1005 A;Molecule type: genomic RNA A;Residues: 776-1127 <YA2> A;Cross-references: UNIPARC:UPI0000174A07 A;Note: the authors translated the codons GTG for residue 899 as Leu, CTG for residue 95 C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: capsid protein; envelope protein; glycoprotein; membrane-associated protein; F;1-114/Product: capsid protein C #status predicted <CAP> F;101-117/Domains: transmembrane #status predicted <TM1> F;115-280/Product: membrane-associated protein M precursor #status predicted <MAM> F;115-205/Domains: nonterminal signal sequence #status predicted <SIG> F;206-280/Product: membrane-associated protein M #status predicted <MEM> F;281-775/Product: envelope protein E #status predicted <ENV> F;727-743/Domains: transmembrane #status predicted <TM2> F;757-773/Domains: transmembrane #status predicted <TM3> F;776-1127/Product: nonstructural protein NS1 #status predicted <NS1> F;183,347,433,905,982/Binding site: carbohydrate (Asn) (covalent) #status predicted	
Query Match 83.6%; Score 46; DB 1; Length 1127; Best Local Similarity 77.8%; Pred. No. 5.7; Mismatches 1; Indels 0; Gaps 0; Matches 7; Conservative 1; Mismatches 0; Gaps 0;	
QY 1 IETWFLRHP 9 :	Db 237 NETWILRHP 245
RESULT 17 GNWVVP	
genome polyprotein - dengue virus type 1 (strain Western Pacific) (fragment) N;Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru C;Species: dengue virus type 1 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 31-Dec-2004 C;Accession: A27032 R;Mason, P.W.; McAda, P.C.; Mason, T.L.; Fournier, M.J. Virology 161, 262-267, 1987 A;Title: Sequence of the dengue-1 virus genome in the region encoding the three structur A;Reference number: A27032; MUID:86044504; PMID:3672932 A;Accession: A27032 A;Molecule type: genomic RNA A;Residues: 1-1226 <MAS> A;Cross-references: UNIPROT:P17763; UNIPARC:UPI0000131DF1; GB:M23027; NID:g511850; PIDN: C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; nucle	

F;2-114/Product: capsid protein C #status predicted <CPC> F;43-59/Domains: transmembrane #status predicted <TM1> F;101-117/Domains: transmembrane #status predicted <TM2> F;115-280/Product: membrane-associated protein M precursor #status predicted <MPP> F;115-205/Domains: nonterminal signal sequence #status predicted <SIG> F;206-280/Product: membrane-associated protein M #status predicted <MPM> F;268-284/Domains: transmembrane #status predicted <TM3> F;281-775/Product: envelope protein E #status predicted <EPE> F;384-391/Region: nucleotide-binding motif A (P-loop) F;727-743/Domains: transmembrane #status predicted <TM4> F;757-773/Domains: transmembrane #status predicted <TM5> F;776-1127/Product: nonstructural protein NS1 #status predicted <NS1> F;1128-1226/Product: nonstructural protein NS2a (fragment) #status predicted <N2A> F;183,347,433,905,982,1190/Binding site: carbohydrate (Asn) (covalent) #status predicted	
Query Match 83.6%; Score 46; DB 1; Length 1226; Best Local Similarity 77.8%; Pred. No. 6.2; Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
QY 1 IETWFLRHP 9 :	Db 237 VETWALRHP 245
RESULT 18 GNWVD3	
genome polyprotein - dengue virus type 3 N;Contains: capsid protein; envelope protein; membrane protein; nonstructural protein NS1 a; nonstructural protein NS4b; nonstructural protein NS5 C;Species: dengue virus type 3 C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004 C;Accession: A34774 R;Oatomi, K.; Sumiyoshi, H. Virology 176, 643-647, 1990 A;Title: Complete nucleotide sequence of dengue type 3 virus genome RNA. A;Reference number: A34774; MUID:90266483; PMID:2345967 A;Accession: A34774 A;Molecule type: genomic RNA A;Residues: 1-3390 <OSA> A;Cross-references: UNIPROT:P27915; UNIPARC:UPI0000131DFE; GB:M93130; NID:g323468; PIDN: C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein; i F;1-114/Product: capsid protein #status predicted <CAP> F;46-67/Domains: transmembrane #status predicted <TM1> F;115-280/Product: membrane protein precursor #status predicted <MEP> F;115-205/Domains: nonterminal signal sequence #status predicted <MEM> F;206-280/Product: membrane protein #status predicted <TM3> F;266-280/Domains: transmembrane #status predicted <TM4> F;281-773/Product: envelope protein #status predicted <ENV> F;724-746/Domains: transmembrane #status predicted <TM5> F;753-771/Domains: transmembrane #status predicted <TM6> F;774-1184/Product: nonstructural protein NS1 #status predicted <NS1> F;1156-1175/Domains: transmembrane #status predicted <TM6> F;1185-1343/Product: nonstructural protein NS2a #status predicted <N2A> F;1344-1473/Product: nonstructural protein NS2b #status predicted <N2B> F;1474-2092/Product: nonstructural protein NS3 #status predicted <NS3> F;1667-1674/Region: nucleotide-binding motif A (P-loop) F;1754-1759/Region: nucleotide-binding motif B F;1758-1761/Region: DEAH motif F;2093-2378/Product: nonstructural protein NS4a #status predicted <N4A> F;2379-2490/Product: nonstructural protein NS4b #status predicted <N4B> F;2491-3390/Product: nonstructural protein NS5 #status predicted <NS5> F;183,347,433,750,903,980,1132,1188,1661,2300,2304,2386,2456,2702,2712/Binding site: car	
Query Match 83.6%; Score 46; DB 1; Length 3390; Best Local Similarity 77.8%; Pred. No. 18; Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
QY 1 IETWFLRHP 9 :	Db 237 VETWALRHP 245

```

RESULT 19
A42551
genome polyprotein - dengue virus type 1 (strain Singapore S275/90)
N:Contains: capsid protein; envelope protein; membrane protein; nonstructural protein NS1
A:Nonstructural protein NS4b; nonstructural protein NS5
C:Species: dengue virus type 1
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004
C:Accession: A42551
R:Fu, J.; Tan, B.H.; Yap, E.H.; Chan, Y.C.; Tan, Y.H.
Virology 188, 953-958, 1992
A:Title: Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90).
A:Reference number: A42551; MUID:92263809; PMID:1585663
A:Accession: A42551
A:Molecule type: genomic RNA
A:Residues: 1-3396 <FU>
A:Cross-references: UNIPROT:P33478; UNIPARC:UPI000002F845; GB:M87512
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
F:1-114/Product: capsid protein #status predicted <CAP>
F:115-281/Product: membrane protein precursor #status predicted <MEP>
F:115-204/Domain: nonterminal signal sequence #status predicted <SIG>
F:205-281/Product: membrane protein #status predicted <MEM>
F:267-279/Domain: transmembrane #status predicted <TM1>
F:282-774/Product: envelope protein #status predicted <ENV>
F:753-769/Domain: transmembrane #status predicted <TM2>
F:775-1127/Product: nonstructural protein NS1 #status predicted <NS1>
F:1128-1344/Product: nonstructural protein NS2a #status predicted <N2A>
F:1345-1474/Product: nonstructural protein NS2b #status predicted <N2B>
F:1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>
F:1688-1675/Region: nucleotide-binding motif A (P-loop)
F:1755-1760/Region: nucleotide-binding motif B
F:1759-1762/Region: DEAH motif
F:2094-2243/Product: nonstructural protein NS4a #status predicted <N4A>
F:2244-2492/Product: nonstructural protein NS4b #status predicted <N4B>
F:2493-3396/Product: nonstructural protein NS5 #status predicted <NS5>
F:183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.6%; Score 46; DB 1; Length 3396;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
: ||| |||
Db 237 VETWALRHP 245

RESULT 20
PS0043
genome polyprotein - dengue virus type 2 (strain PUO-218) (fragment)
N:Contains: envelope protein E; membrane-associated protein M; nonstructural protein NS1
C:Species: dengue virus type 2
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004
C:Accession: PS0043
R:Gruenberg, A.; Woo, W.S.; Biedrzycka, A.; Wright, P.J.
J. Gen. Virol. 69, 1391-1398, 1988
A:Title: Partial nucleotide sequence and deduced amino acid sequence of the structural p
A:Reference number: PS0043; MUID:88258474; PMID:3385407
A:Accession: PS0043
A:Molecule type: mRNA
A:Residues: 1-665 <GRU>
A:Cross-references: UNIPROT:P18356; UNIPARC:UPI0000178550
C:Comment: The RNA sequence was obtained from the DBJ, release 5.0.
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: envelope protein; glycoprotein; membrane protein; nonstructural protein; pol
F:1-31/Domain: signal sequence #status predicted <SIG>
F:192-166/Product: membrane-associated protein M #status predicted <MG>
F:167-661/Product: envelope protein E #status predicted <EPE>
F:662-665/Product: nonstructural protein NS1 (fragment) #status predicted <NS1>
F:69,233,319/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 78.2%; Score 43; DB 2; Length 665;
Best Local Similarity 77.8%; Pred. NO. 11;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 IETWFLRHP 9
: ||| |||
Db 123 IEIWLRLHP 131

RESULT 21
S09224
membrane protein - dengue virus type 2 (strain M2) (fragment)
C:Species: dengue virus type 2
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 31-Dec-2004
C:Accession: S09224
R:Samuel, S.; Koh, C.L.; Pang, T.; Lam, S.K.
Nucleic Acids Res. 18, 1905, 1990
A:Title: Nucleotide and encoded amino acid sequences of the membrane protein precursor a
agic fever, dengue shock syndrome or dengue fever.
A:Reference number: S09223; MUID:90245599; PMID:2336374
A:Accession: S09224
A:Molecule type: genomic RNA
A:Residues: 1-166 <SAM>
A:Cross-references: UNIPROT:Q67422; UNIPARC:UPI00000F4214; EMBL:X51712; NID:G59307; PIDN
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: membrane protein

Query Match 76.4%; Score 42; DB 2; Length 166;
Best Local Similarity 77.8%; Pred. No. 3.9;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
: ||| |||
Db 123 IETWFLRHP 131

RESULT 22
EB6085
hypothetical protein yijF [imported] - Escherichia coli (strain O157:H7, substrain EDJ93)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: EB6085
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: EB6085
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-205 <STO>
A:Cross-references: UNIPROT:Q8X763; UNIPARC:UPI00001659BC; GB:AE005174; NID:G12518859; P
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yijF
C:Superfamily: Escherichia coli hypothetical 23.0K protein b3944

Query Match 74.5%; Score 41; DB 2; Length 205;
Best Local Similarity 75.0%; Pred. No. 7.2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRH 8
: ||| |||
Db 125 LETWFLTRH 132

RESULT 23
A98238
hypothetical protein EC4873 [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: A98238
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

```

A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: A98238  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-205 <HAY>  
A;Cross-references: UNIPROT:Q8X763; UNIPARC:UPI000000D0AB8; GB:BA000007; PIDN:BA038296.1;  
A;Experimental source: strain O157:H7, substrain RMD 0509952  
C;Genetics:  
A;Gene: EC94873  
C;Superfamily: Escherichia coli hypothetical 23.0K protein b3944

Query Match 74.5%; Score 41; DB 2; Length 205;  
Best Local Similarity 75.0%; Pred. No. 7.2; Mismatches 1; Indels 0; Gaps 0;  
Matches 6; Conservative 1

QY 1 IETWFLRH 8  
:|||||  
Db 125 LETWFLRH 132

RESULT 24  
A83184  
Probable protein methyltransferase PA3706 [imported] - Pseudomonas aeruginosa (strain PA  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: A83184  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
Nature 406, 959-964, 2000  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: A83184  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-422 <STO>  
A;Cross-references: UNIPROT:Q9HXT5; UNIPARC:UPI000000C5ACD; GB:AE004789; GB:AE004091; NID  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA3706

Query Match 74.5%; Score 41; DB 2; Length 422;  
Best Local Similarity 75.0%; Pred. No. 15; Mismatches 1; Indels 0; Gaps 0;  
Matches 6; Conservative 1

QY 2 ETWFLRHP 9  
:|||||  
Db 66 ETWFFRYP 73

RESULT 25  
I78665  
hypothetical 23.0K protein b3944 - Escherichia coli (strain K-12)  
N;Alternate names: hypothetical protein F205  
C;Species: Escherichia coli  
C;Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 09-Jul-2004  
C;Accession: I78665; C65201  
R;Blattner, F.R.; Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.  
Nucleic Acids Res. 21, 5408-5417, 1993  
A;Title: Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89  
A;Reference number: 158303; MUID:94089392; PMID:8265357  
A;Accession: I78665  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-205 <RES>  
A;Cross-references: UNIPROT:P32668; UNIPARC:UPI000013B429; EMBL:U000006; NID:9409785; PID  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: C65201  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA  
A;Residues: 1-205 <BLAT>  
A;Cross-references: UNIPARC:UPI000013B429; GB:AE0000468; GB:U00096; NID:G1790374; PIDN:AA  
A;Experimental source: strain K-12, substrain MG1655  
C;Genetics:  
A;Gene: Yijf  
C;Superfamily: Escherichia coli hypothetical 23.0K protein b3944

Query Match 72.7%; Score 40; DB 2; Length 205;  
Best Local Similarity 75.0%; Pred. No. 11; Mismatches 1; Indels 0; Gaps 0;  
Matches 6; Conservative 1

QY 1 IETWFLRH 8  
:|||||  
Db 125 LETWFLRH 132

RESULT 26  
H95879  
Probable sugar ABC transporter permease protein SMB20318 [imported] - Sinorhizobium meli  
C;Species: Sinorhizobium melioli  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: H95879  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: H95879  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-343 <KUR>  
A;Cross-references: UNIPROT:Q92WM8; UNIPARC:UPI00000CB4A7; GB:AL591985; PIDN:CAC48704.1;  
A;Experimental source: strain 1021, megaplasmid pSymb  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont Sinorhizobium melioli.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: SMB20318  
A;Genome: plasmid  
C;Superfamily: l-arabinose transport system permease arah

Query Match 70.9%; Score 39; DB 2; Length 343;  
Best Local Similarity 83.3%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;  
Matches 5; Conservative 1

QY 4 WFLRHP 9  
:|||||  
Db 24 WFLRHP 29

RESULT 27  
GNWVY  
genome polyprotein - yellow fever virus (strain 17D)  
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructura  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: yellow fever virus  
C;Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 31-Dec-2004  
C;Accession: A03914  
R;Rice, C.M.; Leaches, E.M.; Eddy, S.R.; Shin, S.J.; Sheets, R.L.; Strauss, J.H.  
Science 229, 726-733, 1985  
A;Title: Nucleotide sequence of yellow fever virus: implications for flavivirus gene exp  
A;Reference number: A03914; MUID:85272570; PMID:4023707  
A;Accession: A03914  
A;Molecule type: genomic RNA  
A;Residues: 1-3411 <RIC>  
A;Cross-references: UNIPROT:P03314; UNIPARC:UPI0000131E82; GB:X03700; GB:K02749; NID:959  
C;Superfamily: hepatitis C virus genome polyprotein

```

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 IETWFLRHP 9
   |||::||
Db 242 IERNFVRNP 250

RESULT 29
T49934
carboxypeptidase-like protein - Arabidopsis thaliana
N;Alternate names: protein F17114.170
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 05-Oct-2004
R;Accession: T49934
R;Revan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24490
A;Accession: T49934
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-399 <BEV>
A;Cross-references: UNIPROT:O9LXC8; UNIPARC:UPI000009CB29; EMBL:AL353994; GSPDB:GN000063,
A;Experimental source: cultivar Columbia; BAC clone F17114
C;Genetics:
A;Gene: ATSP:F17114.170
A;Map position: 5
A;Antions: 44/2; 93/1; 118/3; 159/3; 194/1; 206/3; 240/3; 269/3; 293/3; 317/2; 356/2
C;Superfamily: Serine carboxypeptidase

Query Match 69.1%; Score 38; DB 2; Length 399;
Best Local Similarity 44.4%; Pred. No. 47;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 IETWFLRHP 9
   :::|||
Db 157 LRSWFVKHP 165

RESULT 30
VGBE14
glycoprotein gpV - human herpesvirus 3
C;Species: human herpesvirus 3, varicella-zoster virus
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004
C;Accession: E27342
R;Davison, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A;Title: The complete DNA sequence of varicella-zoster virus.
A;Reference number: A27345; MUID:86306657; PMID:3018124
A;Accession: E27342
A;Molecule type: DNA
A;Residues: 1-560 <DAV>
A;Cross-references: UNIPROT:P09256; UNIPARC:UPI00001386A3; EMBL:X04370; NID:g59989; PIDN:
C;Genetics:
A;Gene: 14
C;Superfamily: herpesvirus glycoprotein F
C;Keywords: glycoprotein
F;206,325,344,432,461/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.1%; Score 38; DB 1; Length 560;
Best Local Similarity 75.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 IETWFLRH 8
   |||::||
Db 220 IEWFTRH 227

Search completed: August 31, 2006, 11:51:54
Job time : 18.25 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: August 31, 2006, 11:33:43 ; Search time 139 Seconds  
(without alignments)  
59.893 Million cell updates/sec

Title: DENGUE\_SEROTYPE2  
Perfect score: 55  
Sequence: 1 ietwflrhp 9  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2849598 seqs, 925015592 residues  
Total number of hits satisfying chosen parameters: 2849598  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries  
Database : UniProt\_7.2.1\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	89.1	120	2	Q67424 dengue viru
2	49	89.1	166	2	Q66346 dengue viru
3	49	89.1	280	2	Q8QZ64 dengue viru
4	49	89.1	280	2	Q8QZ65 dengue viru
5	49	89.1	280	2	Q8QZ66 dengue viru
6	49	89.1	280	2	Q8QZ67 dengue viru
7	49	89.1	555	1	POLG_DEN2H
8	49	89.1	565	2	Q3ZPJ0 dengue viru
9	49	89.1	578	2	O12290 dengue viru
10	49	89.1	661	2	Q3BCV3 dengue viru
11	49	89.1	661	2	Q3BCV4 dengue viru
12	49	89.1	661	2	Q3BCV5 dengue viru
13	49	89.1	661	2	Q3BCX6 dengue viru
14	49	89.1	661	2	Q3BCX7 dengue viru
15	49	89.1	661	2	Q3BCX8 dengue viru
16	49	89.1	661	2	Q3BCX9 dengue viru
17	49	89.1	661	2	Q3BCY0 dengue viru
18	49	89.1	661	2	Q3BCY1 dengue viru
19	49	89.1	661	2	Q3BCY2 dengue viru
20	49	89.1	661	2	Q3BCY3 dengue viru
21	49	89.1	661	2	Q3BCY4 dengue viru
22	49	89.1	661	2	Q3BCY5 dengue viru
23	49	89.1	661	2	Q5QIB6 dengue viru
24	49	89.1	661	2	Q5VI87 dengue viru
25	49	89.1	661	2	Q5VI88 dengue viru
26	49	89.1	661	2	Q5VI89 dengue viru
27	49	89.1	661	2	Q5VI90 dengue viru
28	49	89.1	661	2	Q5VI91 dengue viru
29	49	89.1	661	2	Q5VI92 dengue viru
30	49	89.1	661	2	Q5VI93 dengue viru
31	49	89.1	661	2	Q5VI94 dengue viru

32	49	89.1	661	2	Q5VI95_9FLAV	Q5vi95 dengue viru
33	49	89.1	661	2	Q5VI96_9FLAV	Q5vi96 dengue viru
34	49	89.1	716	2	Q6DUV2_9FLAV	Q6duv2 dengue viru
35	49	89.1	724	2	Q5ICU9_9FLAV	Q5icu9 dengue viru
36	49	89.1	745	2	Q6KEK9_9FLAV	Q6kek9 dengue viru
37	49	89.1	757	2	Q5S8P1_9FLAV	Q5s8p1 dengue viru
38	49	89.1	757	2	Q5S8P2_9FLAV	Q5s8p2 dengue viru
39	49	89.1	757	2	Q6DUD9_9FLAV	Q6dud9 dengue viru
40	49	89.1	763	2	Q5ICU8_9FLAV	Q5icu8 dengue viru
41	49	89.1	775	2	Q66398_9FLAV	Q66398 dengue viru
42	49	89.1	775	2	Q8QY07_9FLAV	Q8qy07 dengue viru
43	49	89.1	775	2	Q8QY62_9FLAV	Q8qy62 dengue viru
44	49	89.1	775	2	Q8QY63_9FLAV	Q8qy63 dengue viru
45	49	89.1	779	2	Q88636_9FLAV	Q88636 dengue viru
46	49	89.1	1127	2	P87638_9FLAV	P87638 dengue viru
47	49	89.1	1127	2	P89531_9FLAV	P89531 dengue viru
48	49	89.1	1127	2	P89532_9FLAV	P89532 dengue viru
49	49	89.1	1127	2	Q66454_9FLAV	Q66454 dengue viru
50	49	89.1	1127	2	Q66455_9FLAV	Q66455 dengue viru
51	49	89.1	1127	2	Q66456_9FLAV	Q66456 dengue viru
52	49	89.1	1127	2	Q66457_9FLAV	Q66457 dengue viru
53	49	89.1	3388	1	POLG_DEN2P	P12823 d genome po
54	49	89.1	3391	1	POLG_DEN26	P29990 d genome po
55	49	89.1	3391	1	POLG_DEN27	P29991 d genome po
56	49	89.1	3391	1	POLG_DEN2J	P07564 d genome po
57	49	89.1	3391	1	POLG_DEN2N	P14340 d genome po
58	49	89.1	3391	2	O09234_DEN26	O09234 dengue viru
59	49	89.1	3391	2	O11875_9FLAV	O11875 dengue viru
60	49	89.1	3391	2	O92752_9FLAV	O92752 dengue viru
61	49	89.1	3391	2	O92753_9FLAV	O92753 dengue viru
62	49	89.1	3391	2	O92754_9FLAV	O92754 dengue viru
63	49	89.1	3391	2	O92835_9FLAV	O92835 dengue viru
64	49	89.1	3391	2	O58Y66_9FLAV	O58y66 dengue viru
65	49	89.1	3391	2	O58Y67_9FLAV	O58y67 dengue viru
66	49	89.1	3391	2	O58Y69_9FLAV	O58y69 dengue viru
67	49	89.1	3391	2	O58Y71_9FLAV	O58y71 dengue viru
68	49	89.1	3391	2	O5QC63_9FLAV	O5qc63 dengue viru
69	49	89.1	3391	2	Q68Y26_9FLAV	Q68y26 dengue viru
70	49	89.1	3391	2	Q70YQ7_9FLAV	Q70yq7 dengue viru
71	49	89.1	3391	2	Q8QR27_9FLAV	Q8qr27 dengue viru
72	49	89.1	3391	2	Q91SD1_9FLAV	Q91sd1 dengue viru
73	49	89.1	3391	2	Q91U94_9FLAV	Q91u94 dengue viru
74	49	89.1	3391	2	Q9E7P0_9FLAV	Q9e7p0 dengue viru
75	49	89.1	3391	2	Q91F59_9FLAV	Q91f59 dengue viru
76	49	89.1	3391	2	Q9J8D1_9FLAV	Q9j8d1 dengue viru
77	49	89.1	3391	2	Q9J8D2_9FLAV	Q9j8d2 dengue viru
78	49	89.1	3391	2	Q9J8D3_9FLAV	Q9j8d3 dengue viru
79	49	89.1	3391	2	Q9J8D4_9FLAV	Q9j8d4 dengue viru
80	49	89.1	3391	2	Q9J8D5_9FLAV	Q9j8d5 dengue viru
81	49	89.1	3391	2	Q9J8D6_9FLAV	Q9j8d6 dengue viru
82	49	89.1	3391	2	Q9J8D7_9FLAV	Q9j8d7 dengue viru
83	49	89.1	3391	2	Q9J8D8_9FLAV	Q9j8d8 dengue viru
84	49	89.1	3391	2	Q9J8D9_9FLAV	Q9j8d9 dengue viru
85	49	89.1	3391	2	Q9J8E0_9FLAV	Q9j8e0 dengue viru
86	49	89.1	3391	2	Q9J8E1_9FLAV	Q9j8e1 dengue viru
87	49	89.1	3391	2	Q9Q4T1_9FLAV	Q9q4t1 dengue viru
88	49	89.1	3391	2	Q9Q4T2_9FLAV	Q9q4t2 dengue viru
89	49	89.1	3391	2	Q9W8I3_9FLAV	Q9w8i3 dengue viru
90	49	89.1	3391	2	Q9WD99_9FLAV	Q9wd99 dengue viru
91	49	89.1	3391	2	Q9WDA0_9FLAV	Q9wda0 dengue viru
92	49	89.1	3391	2	Q9WDA1_9FLAV	Q9wda1 dengue viru
93	49	89.1	3391	2	Q9WDA2_9FLAV	Q9wda2 dengue viru
94	49	89.1	3391	2	Q9WDA3_9FLAV	Q9wda3 dengue viru
95	49	89.1	3391	2	Q9WDA4_9FLAV	Q9wda4 dengue viru
96	49	89.1	3391	2	Q9WDA5_9FLAV	Q9wda5 dengue viru
97	49	89.1	3391	2	Q9WDA6_9FLAV	Q9wda6 dengue viru
98	49	89.1	3391	2	Q9WDA7_9FLAV	Q9wda7 dengue viru
99	49	89.1	3391	2	Q9WL24_9FLAV	Q9wl24 dengue viru
100	49	89.1	3391	2	Q9WL25_9FLAV	Q9wl25 dengue viru

ALIGNMENTS

```
-----
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      EMBL; X72849; CAA51363.1; -; mRNA.
DR      PIR; S40144; S40144.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019058; P:viral infectious cycle; IEA.
DR      InterPro; IPR000069; Flavi_M.
DR      InterPro; IPR002535; Flavi_propep.
DR      Pfam; PF01004; Flavi_M; 1.
DR      Pfam; PF01570; Flavi_propep; 1.
KW      Polyprotein.
FT      CHAIN          92 >166      membrane protein.
FT      NON_TER       1          1
FT      NON_TER      166      166
FT      NON_TER      166      166
SQ      SEQUENCE 166 AA; 18751 MW; F498748D35909639 CRC64;

Query Match      89.1%; Score 49; DB 2; Length 166;
Best Local Similarity 88.9%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IETWFLRHP 9
DB      123 IETWILRHP 131

RESULT 3
Q8QZ64_9FLAV PRELIMINARY; PRT; 280 AA.
AC      Q8QZ64;
DT      01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT      01-JUN-2002, sequence version 1.
DT      07-FEB-2006, entry version 9.
DE      Polyprotein (Fragment).
OS      Dengue virus type 2.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Flavivirus; Dengue virus group.
OX      NCBI_TaxID=11060;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=21571640; PubMed=11714970;
RA      Holmes E.C., Gould E.A.;
RA      "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT      for in situ virus evolution and recombination.";
RL      J. Gen. Virol. 82:2945-2953(2001).
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      EMBL; AF360863; AAL76291.1; -; Genomic_RNA.
DR      SMR; Q8QZ64; 21-100.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0005198; P:structural molecule activity; IEA.
DR      GO; GO:0019058; P:viral infectious cycle; IEA.
DR      InterPro; IPR001122; Flavi_capsidC.
DR      InterPro; IPR000069; Flavi_M.
DR      InterPro; IPR002535; Flavi_propep.
DR      Pfam; PF01003; Flavi_capsid; 1.
DR      Pfam; PF01004; Flavi_M; 1.
DR      Pfam; PF01570; Flavi_propep; 1.
KW      Polyprotein.
FT      NON_TER       280      280
FT      NON_TER      31847 MW; E889FDD11929CBA7 CRC64;
SQ      SEQUENCE 280 AA; 31847 MW; E889FDD11929CBA7 CRC64;

Query Match      89.1%; Score 49; DB 2; Length 280;
Best Local Similarity 88.9%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IETWFLRHP 9
DB      237 IETWILRHP 245

-----
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      EMBL; X05375; CAA28966.1; -; Genomic_RNA.
DR      HSP; Q88653; 10KE.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019031; C:viral envelope; IEA.
DR      GO; GO:0019058; P:viral infectious cycle; IEA.
DR      InterPro; IPR011999; Flavi_glyc_cen_dm.
DR      InterPro; IPR000069; Flavi_M.
DR      InterPro; IPR011998; v1 glye cen dim.
DR      Pfam; PF00869; Flavi_glycoprot; 1.
DR      Pfam; PF01004; Flavi_M; 1.
KW      Envelope protein.
FT      CHAIN          18      92      protein M.
FT      CHAIN          93 >120      protein E.
FT      NON_TER       1          1
FT      NON_TER      120      120
FT      NON_TER      120      120
SQ      SEQUENCE 120 AA; 13329 MW; FF86913787CA5C27 CRC64;

Query Match      89.1%; Score 49; DB 2; Length 120;
Best Local Similarity 88.9%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IETWFLRHP 9
DB      49 IETWILRHP 57

RESULT 2
Q66346_9FLAV PRELIMINARY; PRT; 166 AA.
AC      Q66346;
DT      01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT      01-NOV-1996, sequence version 1.
DT      07-FEB-2006, entry version 21.
DE      Premembrane polyprotein (Fragment).
OS      Dengue virus type 2.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Flavivirus; Dengue virus group.
OX      NCBI_TaxID=11060;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      STRAIN=TH-36;
RA      Shiu S.Y.W.;
RA      Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      EMBL; X72849; CAA51363.1; -; mRNA.
DR      PIR; S40144; S40144.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019058; P:viral infectious cycle; IEA.
DR      InterPro; IPR000069; Flavi_M.
DR      InterPro; IPR002535; Flavi_propep.
DR      Pfam; PF01004; Flavi_M; 1.
DR      Pfam; PF01570; Flavi_propep; 1.
KW      Polyprotein.
FT      CHAIN          92 >166      membrane protein.
FT      NON_TER       1          1
FT      NON_TER      166      166
FT      NON_TER      166      166
SQ      SEQUENCE 166 AA; 18751 MW; F498748D35909639 CRC64;

Query Match      89.1%; Score 49; DB 2; Length 166;
Best Local Similarity 88.9%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IETWFLRHP 9
DB      123 IETWILRHP 131

RESULT 3
Q8QZ64_9FLAV PRELIMINARY; PRT; 280 AA.
AC      Q8QZ64;
DT      01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT      01-JUN-2002, sequence version 1.
DT      07-FEB-2006, entry version 9.
DE      Polyprotein (Fragment).
OS      Dengue virus type 2.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Flavivirus; Dengue virus group.
OX      NCBI_TaxID=11060;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=21571640; PubMed=11714970;
RA      Holmes E.C., Gould E.A.;
RA      "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT      for in situ virus evolution and recombination.";
RL      J. Gen. Virol. 82:2945-2953(2001).
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      EMBL; AF360863; AAL76291.1; -; Genomic_RNA.
DR      SMR; Q8QZ64; 21-100.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0005198; P:structural molecule activity; IEA.
DR      GO; GO:0019058; P:viral infectious cycle; IEA.
DR      InterPro; IPR001122; Flavi_capsidC.
DR      InterPro; IPR000069; Flavi_M.
DR      InterPro; IPR002535; Flavi_propep.
DR      Pfam; PF01003; Flavi_capsid; 1.
DR      Pfam; PF01004; Flavi_M; 1.
DR      Pfam; PF01570; Flavi_propep; 1.
KW      Polyprotein.
FT      NON_TER       280      280
FT      NON_TER      31847 MW; E889FDD11929CBA7 CRC64;
SQ      SEQUENCE 280 AA; 31847 MW; E889FDD11929CBA7 CRC64;

Query Match      89.1%; Score 49; DB 2; Length 280;
Best Local Similarity 88.9%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IETWFLRHP 9
DB      237 IETWILRHP 245

-----
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      EMBL; X05375; CAA28966.1; -; Genomic_RNA.
DR      HSP; Q88653; 10KE.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019031; C:viral envelope; IEA.
DR      GO; GO:0019058; P:viral infectious cycle; IEA.
DR      InterPro; IPR011999; Flavi_glyc_cen_dm.
DR      InterPro; IPR000069; Flavi_M.
DR      InterPro; IPR011998; v1 glye cen dim.
DR      Pfam; PF00869; Flavi_glycoprot; 1.
DR      Pfam; PF01004; Flavi_M; 1.
KW      Envelope protein.
FT      CHAIN          18      92      protein M.
FT      CHAIN          93 >120      protein E.
FT      NON_TER       1          1
FT      NON_TER      120      120
FT      NON_TER      120      120
SQ      SEQUENCE 120 AA; 13329 MW; FF86913787CA5C27 CRC64;

Query Match      89.1%; Score 49; DB 2; Length 120;
Best Local Similarity 88.9%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IETWFLRHP 9
DB      49 IETWILRHP 57

RESULT 2
Q66346_9FLAV PRELIMINARY; PRT; 166 AA.
AC      Q66346;
DT      01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT      01-NOV-1996, sequence version 1.
DT      07-FEB-2006, entry version 21.
DE      Premembrane polyprotein (Fragment).
OS      Dengue virus type 2.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Flavivirus; Dengue virus group.
OX      NCBI_TaxID=11060;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      STRAIN=TH-36;
RA      Shiu S.Y.W.;
RA      Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      EMBL; X72849; CAA51363.1; -; mRNA.
DR      PIR; S40144; S40144.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019058; P:viral infectious cycle; IEA.
DR      InterPro; IPR000069; Flavi_M.
DR      InterPro; IPR002535; Flavi_propep.
DR      Pfam; PF01004; Flavi_M; 1.
DR      Pfam; PF01570; Flavi_propep; 1.
KW      Polyprotein.
FT      CHAIN          92 >166      membrane protein.
FT      NON_TER       1          1
FT      NON_TER      166      166
FT      NON_TER      166      166
SQ      SEQUENCE 166 AA; 18751 MW; F498748D35909639 CRC64;

Query Match      89.1%; Score 49; DB 2; Length 166;
Best Local Similarity 88.9%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IETWFLRHP 9
DB      123 IETWILRHP 131

RESULT 3
Q8QZ64_9FLAV PRELIMINARY; PRT; 280 AA.
AC      Q8QZ64;
DT      01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT      01-JUN-2002, sequence version 1.
DT      07-FEB-2006, entry version 9.
DE      Polyprotein (Fragment).
OS      Dengue virus type 2.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Flavivirus; Dengue virus group.
OX      NCBI_TaxID=11060;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=21571640; PubMed=11714970;
RA      Holmes E.C., Gould E.A.;
RA      "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT      for in situ virus evolution and recombination.";
RL      J. Gen. Virol. 82:2945-2953(2001).
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      EMBL; AF360863; AAL76291.1; -; Genomic_RNA.
DR      SMR; Q8QZ64; 21-100.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0005198; P:structural molecule activity; IEA.
DR      GO; GO:0019058; P:viral infectious cycle; IEA.
DR      InterPro; IPR001122; Flavi_capsidC.
DR      InterPro; IPR000069; Flavi_M.
DR      InterPro; IPR002535; Flavi_propep.
DR      Pfam; PF01003; Flavi_capsid; 1.
DR      Pfam; PF01004; Flavi_M; 1.
DR      Pfam; PF01570; Flavi_propep; 1.
KW      Polyprotein.
FT      NON_TER       280      280
FT      NON_TER      31847 MW; E889FDD11929CBA7 CRC64;
SQ      SEQUENCE 280 AA; 31847 MW; E889FDD11929CBA7 CRC64;

Query Match      89.1%; Score 49; DB 2; Length 280;
Best Local Similarity 88.9%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 IETWFLRHP 9
DB      237 IETWILRHP 245

-----
```



DR SMR; Q8QZ66; 21-100.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0019058; P:viral infectious cycle; IEA.  
DR InterPro; IPR001122; Flavi\_capsidC.  
DR InterPro; IPR000069; Flavi\_M.  
DR InterPro; IPR002535; Flavi\_propep.  
DR Pfam; PF01003; Flavi\_capsid; 1.  
DR Pfam; PF01004; Flavi\_M; 1.  
DR Pfam; PF01570; Flavi\_propep; 1.  
KW Polyprotein.  
FT NON\_TER 280 280  
SQ SEQUENCE 280 AA; 31893 MW; 814A8B9B422AC20 CRC64;  
Query Match 89.1%; Score 49; DB 2; Length 280;  
Best Local Similarity 88.9%; Pred.No. 3.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps  
QY 1 IETWFLRHP 9  
DB 237 IETWILRHP 245  
RESULT 6  
Q8QZ67\_9FLAV PRELIMINARY; PRT; 280 AA.  
ID Q8QZ67\_9FLAV  
AC Q8QZ67;  
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.  
DT 01-JUN-2002, sequence version 1.  
DT 07-FEB-2006, entry version 9.  
DS Polyprotein (Fragment).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; dengue virus group.  
OX NCBI\_TaxID=11060;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=21571640; PubMed=11714970;  
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,  
RA Holmes E.C., Gould E.A.;  
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence  
RT for in situ virus evolution and recombination.";  
RL J. Gen. Virol. 82:2945-2953(2001).  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
CC EMBL; AF360860; AAL76288.1; -; Genomic\_RNA.  
DR SMR; Q8QZ67; 21-100.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0019058; P:viral infectious cycle; IEA.  
DR InterPro; IPR001122; Flavi\_capsidC.  
DR InterPro; IPR000069; Flavi\_M.  
DR InterPro; IPR002535; Flavi\_propep.  
DR Pfam; PF01003; Flavi\_capsid; 1.  
DR Pfam; PF01004; Flavi\_M; 1.  
DR Pfam; PF01570; Flavi\_propep; 1.  
KW Polyprotein.  
FT NON\_TER 280 280  
SQ SEQUENCE 280 AA; 31893 MW; 814A8B9B422AC20 CRC64;  
Query Match 89.1%; Score 49; DB 2; Length 280;  
Best Local Similarity 88.9%; Pred.No. 3.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps  
QY 1 IETWFLRHP 9  
DB 237 IETWILRHP 245  
RESULT 7  
POLG DEN2H





```
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SLMC70/1995/human;
RX PubMed=16222028;
RA Salda L.T.D.;
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SLMC70/1995/human;
RA Salda L.T.D.;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786396; AAX18214.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73081 MW; 6F6C51D6BEC33CA8 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 13
Q3BCX6_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX6;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St23/2000;
RX PubMed=16222028;
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St23/2000;
RX PubMed=16222028;
RA Salda L.T.;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786375; AAX18193.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73081 MW; 5F56106DA1550EF6 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 14
Q3BCX7_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX7;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St22/2000;
RX PubMed=16222028;
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St22/2000;
RX PubMed=16222028;
RA Salda L.T.D.;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786374; AAX18192.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73083 MW; 5F56106DA1550EF6 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 15
Q3BCX8_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX8;
DR EMBL; AY786375; AAX18193.1; -; Genomic RNA.
```

```
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73143 MW; 5509B5931AE2BF2B CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 14
Q3BCX7_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX7;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St22/2000;
RX PubMed=16222028;
RA Salda L.T.; Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St22/2000;
RX PubMed=16222028;
RA Salda L.T.D.;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786374; AAX18192.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73083 MW; 5F56106DA1550EF6 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 15
Q3BCX8_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX8;
DR EMBL; AY786375; AAX18193.1; -; Genomic RNA.
```



```
RESULT 18
Q3BCY1_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCY1_9FLAV
AC Q3BCY1;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CSMC7/1996;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CSMC7/1996;
RX PubMed=16222028;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL: AV786370; AAX18188.1; -; Genomic RNA.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR GO: 0019058; P: viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 661
FT NON_TER 661
SQ SEQUENCE 661 AA; 73054 MW; 751344A7E73C46F CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 19
Q3BCY2_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCY2_9FLAV
AC Q3BCY2;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRL3/1996;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRL3/1996;
RX PubMed=16222028;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL: AV786368; AAX18186.1; -; Genomic RNA.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR GO: 0019058; P: viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 661
FT NON_TER 661
SQ SEQUENCE 661 AA; 73096 MW; CF865AAE54ADE0F1 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 20
Q3BCY3_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCY3_9FLAV
AC Q3BCY3;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH97/1995;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH97/1995;
RX PubMed=16222028;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL: AV786369; AAX18187.1; -; Genomic RNA.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR GO: 0019058; P: viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 661
FT NON_TER 661
SQ SEQUENCE 661 AA; 73072 MW; 654A28D6B96FB5A0 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131
```

```

Db      123 IETWLRHP 131
-----
RESULT 21
Q3BCY4_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCY4_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY4;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH90/1995;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH90/1995;
RA Salda L.T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY786367; AAX18185.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 661
FT NON_TER 661
SQ SEQUENCE 661 AA; 73150 MW; 654F225FA969639F CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 IETWLRHP 9
      123 IETWLRHP 131
-----
RESULT 22
Q3BCY5_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCY5_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY5;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SLMC125/1995;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SLMC125/1995;
RA Salda L.T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY786367; AAX18185.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 661
FT NON_TER 661
SQ SEQUENCE 661 AA; 73150 MW; 654F225FA969639F CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 IETWLRHP 9
      123 IETWLRHP 131
-----

```

```

RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SLMC125/1995;
RA Salda L.T.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY786366; AAX18184.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1 661
FT NON_TER 661
SQ SEQUENCE 661 AA; 73072 MW; 654A28D6B96FB5A8 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 IETWLRHP 9
      123 IETWLRHP 131
-----
RESULT 23
Q5QIB6_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q5QIB6_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q5QIB6;
DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
DT 04-JAN-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BC134-Merida-94;
RX PubMed=15516647;
RA Llorca-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY466449; AAS45234.1; -; mRNA.
DR SMR; Q5QIB6; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.

```

```

DR Pfam: PF01570; Flavi_propep; 1.
KW Polyprotein. 1 1
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73207 MW; A919612986E04157 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 24
QSVI87_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI87;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=13382/Tizimin 02;
RX PubMed=15516647;
RA Loroño-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AY449684; AAS14975.1; -; Genomic_RNA.
DR SMR; QSVI87; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycop_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 661 661
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73119 MW; CE2051C17F40A623 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 25
QSVI87_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI87;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=13382/Tizimin 02;
RX PubMed=15516647;
RA Loroño-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AY449684; AAS14975.1; -; Genomic_RNA.
DR SMR; QSVI87; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycop_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 661 661
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73119 MW; CE2051C17F40A623 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 25

```

```

QSVI88_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI88;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=13381/Chochola 02;
RX PubMed=15516647;
RA Loroño-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AY449683; AAS14974.1; -; Genomic_RNA.
DR SMR; QSVI88; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycop_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 661 661
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73092 MW; 482C14A6B3B179FA CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 26
QSVI89_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI89;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=12021/Oxkutzcab 01;
RX PubMed=15516647;
RA Loroño-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,

```



```

RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RL the Yucatan State of Mexico.";
CC Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY449682; AAS14973.1; -; Genomic_RNA.
DR SMR; QSV191; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; F:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF02832; Flavi_glyc_C; 1.
DR Pfam; PF00869; Flavi_glyc_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
DR Polyprotein.
DR NON_TER 1
DR NON_TER 661
DR SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 27
QSV190_9FLAV PRELIMINARY; PRT; 661 AA.
ID QSV190_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSV190;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=11936/St. Elena 01;
RX PubMed=15516647;
RA Loroño-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RL the Yucatan State of Mexico.";
CC Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY449681; AAS14972.1; -; Genomic_RNA.
DR SMR; QSV190; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF02832; Flavi_glyc_C; 1.
DR Pfam; PF00869; Flavi_glyc_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
DR Polyprotein.
DR NON_TER 1
DR NON_TER 661
DR SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9
Db 123 IETWILRHP 131

RESULT 28
QSV191_9FLAV PRELIMINARY; PRT; 661 AA.
ID QSV191_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSV191;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=12914/Tekax 01;
RX PubMed=15516647;
RA Loroño-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RL the Yucatan State of Mexico.";
CC Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY449680; AAS14971.1; -; Genomic_RNA.
DR SMR; QSV191; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; F:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF02832; Flavi_glyc_C; 1.
DR Pfam; PF00869; Flavi_glyc_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
DR Polyprotein.
DR NON_TER 1
DR NON_TER 661
DR SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;
Best Local Similarity 88.9%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IETWFLRHP 9
Db 123 IETWILRHP 131

```

Best Local Similarity 88.9%; Pred. No. 7.4; Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9  
|||||  
Db 123 IETWILRHP 131

## RESULT 29

QSVI92\_9FLAV PRELIMINARY; PRT; 661 AA.  
ID QSVI92;  
AC QSVI92;  
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 07-DEC-2004, sequence version 1.  
DT 07-FEB-2006, entry version 8.  
DE Polyprotein (Fragment).  
OS Dengue virus type 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; Dengue virus group.  
OX NCBI\_TaxID=11060;  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C1077/Chilpancingo 97;  
RX PubMed=15516647;  
RA Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L., Diaz F.J.,  
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Blair C.D., Olson K.E., Black W. IV,  
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,  
RA Beaty B.J.;  
RT "Introduction of the American/Asian genotype of dengue 2 virus into  
the Yucatan State of Mexico.";  
RL Am. J. Trop. Med. Hyg. 71:485-492 (2004).  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
-----  
EMBL: AY449679; AAS14970.1; -; Genomic\_RNA.  
DR SMR; QSVI92; 167-560.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; P:structural molecule activity; IEA.  
DR GO; GO:0019058; P:viral infectious cycle; IEA.  
DR InterPro; IPR011999; Flav\_glyc\_cen\_dm.  
DR InterPro; IPR000069; Flavi\_M.  
DR InterPro; IPR002535; Flavi\_propep.  
DR InterPro; IPR000336; Flv\_glyc\_ig-like.  
DR InterPro; IPR011998; Vrl\_glyc\_cen\_dim.  
DR Pfam; PF02832; Flavi\_glycop\_C1.  
DR Pfam; PF00869; Flavi\_glycoprot; 1.  
DR Pfam; PF01004; Flavi\_M; 1.  
DR Pfam; PF01570; Flavi\_propep; 1.  
KW Polyprotein.  
FT NON\_TER 1  
FT NON\_TER 661  
SQ SEQUENCE 661 AA; 73024 MW; 367D6F1A9F25932B CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;  
Best Local Similarity 88.9%; Pred. No. 7.4;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9  
|||||  
Db 123 IETWILRHP 131

## RESULT 30

QSVI93\_9FLAV PRELIMINARY; PRT; 661 AA.  
ID QSVI93;  
AC QSVI93;  
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.  
DT 07-DEC-2004, sequence version 1.  
DT 07-FEB-2006, entry version 8.  
DE Polyprotein (Fragment).

Dengue virus type 2.  
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus; Dengue virus group.  
OX NCBI\_TaxID=11060;  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C-932/Acapulco 97;  
RX PubMed=15516647;  
RA Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,  
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,  
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,  
RA Beaty B.J.;  
RT "Introduction of the American/Asian genotype of dengue 2 virus into  
the Yucatan State of Mexico.";  
RL Am. J. Trop. Med. Hyg. 71:485-492 (2004).  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
-----  
EMBL: AY449678; AAS14969.1; -; Genomic\_RNA.  
DR SMR; QSVI93; 167-560.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005198; P:structural molecule activity; IEA.  
DR GO; GO:0019058; P:viral infectious cycle; IEA.  
DR InterPro; IPR011999; Flav\_glyc\_cen\_dm.  
DR InterPro; IPR000069; Flavi\_M.  
DR InterPro; IPR002535; Flavi\_propep.  
DR InterPro; IPR000336; Flv\_glyc\_ig-like.  
DR InterPro; IPR011998; Vrl\_glyc\_cen\_dim.  
DR Pfam; PF02832; Flavi\_glycop\_C1.  
DR Pfam; PF00869; Flavi\_glycoprot; 1.  
DR Pfam; PF01004; Flavi\_M; 1.  
DR Pfam; PF01570; Flavi\_propep; 1.  
KW Polyprotein.  
FT NON\_TER 1  
FT NON\_TER 661  
SQ SEQUENCE 661 AA; 73024 MW; 0E74A2AC438791A1 CRC64;

Query Match 89.1%; Score 49; DB 2; Length 661;  
Best Local Similarity 88.9%; Pred. No. 7.4;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IETWFLRHP 9  
|||||  
Db 123 IETWILRHP 131

Search completed: August 31, 2006, 11:43:10  
Job time : 140 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: August 31, 2006, 11:33:43 ; Search time 110.25 Seconds  
(without alignments)  
37.324 Million cell updates/sec

Title: DENGUE\_SEROTYPE3  
Perfect score: 56  
Sequence: 1 retwflrhp 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues  
Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_Geneseq\_8.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*  
10: Geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length DB ID	Description			
1	51	91.1	39	9	ADW12582	-----	Adw12582 M1-40/DEN
2	51	91.1	48	9	ADM12588	-----	Adw12588 p(95-114)
3	45	80.4	9	9	ADM12595	-----	Adw12595 M32-40/DE
4	45	80.4	21	9	ADM12594	-----	Adw12594 M20-40/DE
5	45	80.4	32	9	ADM12593	-----	Adw12593 M10-40/DE
6	45	80.4	39	9	ADM12576	-----	Adw12576 M1-40/DEN
7	45	80.4	40	5	AAE17432	-----	AAE17432 Dengue (D
8	45	80.4	40	9	ADM12578	-----	Adw12578 M1-40/YF
9	45	80.4	48	5	AAE17433	-----	AAE17433 (95-114)E
10	45	80.4	167	8	ADN37497	-----	Adn37497 Dengue vi
11	45	80.4	171	8	ADN37493	-----	Adn37493 Dengue vi
12	45	80.4	171	8	ADN37496	-----	Adn37496 Dengue vi
13	45	80.4	635	2	AAW75410	-----	AAW75410 Fusion pr
14	45	80.4	675	8	ADN37628	-----	Adn37628 Dengue vi
15	45	80.4	675	8	ADN37518	-----	Adn37518 Dengue vi
16	45	80.4	675	8	ADN37612	-----	Adn37612 Dengue vi
17	45	80.4	677	2	AAW75411	-----	AAW75411 Fusion pr
18	45	80.4	677	2	AAW75411	-----	AAW75411 Fusion pr
19	45	80.4	677	8	ADN37613	-----	Adn37613 Dengue vi
20	45	80.4	681	8	ADN37603	-----	Adn37603 Dengue vi
21	45	80.4	681	8	ADN37517	-----	Adn37517 Dengue vi
22	45	80.4	685	6	ABP57874	-----	Abp57874 Plasmid p
23	45	80.4	685	6	ABP57876	-----	Abp57876 Plasmid p

24	45	80.4	685	6	ABP57875	-----	Abp57875 Plasmid p
25	45	80.4	1127	2	AAW09409	-----	Aaw09409 Dengue vi
26	45	80.4	1127	2	AAW05522	-----	Aaw05522 Dengue vi
27	45	80.4	1127	7	ADL98086	-----	Adl98086 Dengue vi
28	45	80.4	1127	8	ADQ28716	-----	Adq28716 Dengue vi
29	45	80.4	3388	6	AAE35314	-----	AAE35314 Dengue vi
30	45	80.4	3391	2	AAW06591	-----	Aaw06591 Polypeptide
31	45	80.4	3391	2	AAW06591	-----	Aaw06591 Polypeptide
32	45	80.4	3391	2	AAW06590	-----	Aaw06590 Polypeptide
33	45	80.4	3391	4	AAE07987	-----	Aae07987 Attenuate
34	45	80.4	3391	4	AAE07986	-----	Aae07986 Wild-type
35	45	80.4	3391	8	ADG93314	-----	Adg93314 DEN2 (Ton
36	43	76.8	9	9	ADM12597	-----	Adm12597 M32-40/DE
37	43	76.8	40	5	AAE17431	-----	AAE17431 Dengue (D
38	43	76.8	48	5	AAE17437	-----	AAE17437 (95-114)E
39	43	76.8	55	5	AAE17438	-----	AAE17438 p(95-114)E
40	43	76.8	167	8	ADN37494	-----	Adn37494 Dengue vi
41	43	76.8	167	8	ADN37501	-----	Adn37501 Dengue vi
42	43	76.8	167	8	ADN37498	-----	Adn37498 Dengue vi
43	43	76.8	167	8	ADN37492	-----	Adn37492 Dengue vi
44	43	76.8	167	8	ADN37500	-----	Adn37500 Dengue vi
45	43	76.8	675	8	ADN37624	-----	Adn37624 Dengue vi
46	43	76.8	675	8	ADN37519	-----	Adn37519 Dengue vi
47	43	76.8	675	8	ADN37521	-----	Adn37521 Dengue vi
48	43	76.8	675	8	ADN37616	-----	Adn37616 Dengue vi
49	43	76.8	675	8	ADN37523	-----	Adn37523 Dengue vi
50	43	76.8	675	8	ADN37621	-----	Adn37621 Dengue vi
51	43	76.8	675	8	ADN37604	-----	Adn37604 Dengue vi
52	43	76.8	675	8	ADN37614	-----	Adn37614 Dengue vi
53	43	76.8	675	8	ADN37618	-----	Adn37618 Dengue vi
54	43	76.8	675	8	ADN37620	-----	Adn37620 Dengue vi
55	43	76.8	675	8	ADN37615	-----	Adn37615 Dengue vi
56	43	76.8	675	8	ADN37611	-----	Adn37611 Dengue vi
57	43	76.8	675	8	ADN37619	-----	Adn37619 Dengue vi
58	43	76.8	677	8	ADN37617	-----	Adn37617 Dengue vi
59	43	76.8	677	8	ADN37522	-----	Adn37522 DEN-1/ben
60	43	76.8	677	8	ADN37602	-----	Adn37602 Dengue vi
61	43	76.8	677	8	ADN37515	-----	Adn37515 Dengue vi
62	43	76.8	679	8	ADS76179	-----	Adg76179 Heterodim
63	43	76.8	681	8	ADN37622	-----	Adn37622 Dengue vi
64	43	76.8	684	8	ADR87180	-----	Adr87180 Dengue vi
65	43	76.8	715	2	AAW06593	-----	Aaw06593 Amino aci
66	43	76.8	774	8	ADG93320	-----	Adg93320 DEN1 (Pue
67	43	76.8	775	8	ADG93318	-----	Adg93318 DEN1 (Pue
68	43	76.8	798	2	AAW06592	-----	Aaw06592 Amino aci
69	43	76.8	3389	4	AAE07984	-----	Aae07984 Dengue vi
70	43	76.8	3390	4	AAE07989	-----	Aae07989 Wild-type
71	43	76.8	3390	4	AAE07990	-----	Aae07990 Attenuate
72	43	76.8	3391	4	AAE07982	-----	Aae07982 Dengue vi
73	43	76.8	3391	4	AAE07983	-----	Aae07983 Dengue vi
74	43	76.8	3391	4	AAE07993	-----	Aae07993 Dengue vi
75	43	76.8	3392	4	AAE07981	-----	Aae07981 Attenuate
76	43	76.8	3392	4	AAE07980	-----	Aae07980 Wild-type
77	43	76.8	3396	2	AAR43662	-----	Aar43662 DEN1-S275
78	42	75.0	278	5	ADQ25888	-----	Adq25888 Human GPC
79	42	75.0	826	5	ABB07253	-----	Abb07253 Human nov
80	42	75.0	827	6	ABU075568	-----	Abu075568 Human sec
81	42	75.0	904	4	ABG09947	-----	Abg09947 Novel hum
82	42	75.0	924	5	AAW71323	-----	Aaw71323 Human GCR
83	42	75.0	953	7	ADE34415	-----	Ade34415 Human G-p
84	42	75.0	994	5	ABB07252	-----	Abb07252 Human nov
85	42	75.0	994	5	AAU99808	-----	Aau99808 Novel hum
86	42	75.0	994	7	ADE34425	-----	Ade34425 Human G-p
87	42	75.0	994	8	ADO28977	-----	Ado28977 Human nov
88	42	75.0	994	8	ADQ25892	-----	Adq25892 Human gua
89	42	75.0	1018	5	AAE25061	-----	Aae25061 Human G-p
90	42	75.0	1070	6	ABU07567	-----	Abu07567 Human sec
91	42	75.0	1131	4	ABG11655	-----	Abg11655 Novel hum
92	42	75.0	1232	7	ADF70474	-----	Adf70474 Orphan re
93	41	73.2	39	9	ADM12599	-----	Adm12599 M1-40/DEN
94	41	73.2	455	7	ABO69519	-----	Aboc69519 Pseudomon
95	40	71.4	364	8	ADX92172	-----	Adx92172 Plant ful
96	40	71.4	611	6	ABU22889	-----	Abu22889 Protein e

97 39 69.6 150 1 AAP91166 Aap91166 PUO-218 s  
 98 39 69.6 161 3 AAG03970 Aag03970 Human sec  
 99 39 69.6 203 8 ADY10913 Ady10913 Plant ful  
 100 39 69.6 661 4 AAB84901 Aab84901 Dengue-2

## ALIGNMENTS

## RESULT 1

ADW12582  
 ID ADW12582 standard; peptide; 39 AA.  
 XX  
 AC ADW12582;  
 DT 24-MAR-2005 (first entry)  
 XX  
 DE M1-40/DEN-2 (F36) mutant protein.  
 XX  
 KW Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 KW DEN; dengue; mutant; mutein.  
 XX  
 OS Dengue virus.  
 XX  
 PN US2004266987-A1.  
 XX  
 PD 30-DEC-2004.  
 XX  
 PF 30-JUN-2003; 2003US-00608029.  
 XX  
 PR 30-JUN-2003; 2003US-00608029.  
 XX  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Despres P, Catteau A;  
 XX  
 DR WPI; 2005-047647/05.  
 XX  
 XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
 PT as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.  
 XX  
 PS Example 1; SEQ ID NO 29; 30pp; English.  
 XX  
 CC The present invention relates to an isolated and purified ApoptoM  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC M1-40/DEN (dengue)-2 (F36) mutant protein.  
 XX  
 SQ Sequence 39 AA;

Query Match 91.1%; Score 51; DB 9; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 0.088;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY .2 ETWFLRHP 9  
 DB 32 ETWFLRHP 39  
 |||||  
 RESULT 2  
 ADW12588  
 ID ADW12588 standard; protein; 48 AA.  
 XX  
 AC ADW12588;  
 XX  
 DT 24-MAR-2005 (first entry)  
 KW Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 KW DEN; dengue; mutant; mutein.

DE p(95-114) EGFP (M1-M40) DEN-2 (136F) plasmid DNA encoded protein #3.  
 XX  
 KW Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 KW DEN; dengue; EGFP; enhanced green fluorescent protein.  
 XX  
 OS Dengue virus.  
 OS Chimeric.  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2 /note= "Encoded by GGC"  
 FT Misc-difference 4 /note= "Encoded by GAC"  
 FT Misc-difference 13..44 /note= "Encoded by GTTTC"  
 FT  
 XX  
 PN US2004266987-A1.  
 XX  
 PD 30-DEC-2004.  
 XX  
 PF 30-JUN-2003; 2003US-00608029.  
 XX  
 PR 30-JUN-2003; 2003US-00608029.  
 XX  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Despres P, Catteau A;  
 XX  
 DR WPI; 2005-047647/05.  
 DR N-PSDB; ADW12589.  
 XX  
 XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful  
 PT as a vaccine for preventing or treating pathological conditions from non-  
 PT specific febrile illnesses to severe hemorrhagic manifestations or  
 PT encephalitic syndromes.  
 XX  
 PS Disclosure; SEQ ID NO 35; 30pp; English.  
 XX  
 CC The present invention relates to an isolated and purified ApoptoM  
 CC peptide. The invention is useful as a vaccine for the prevention and  
 CC treatment of pathological conditions from non-specific febrile illnesses  
 CC to severe hemorrhagic manifestations, encephalitic syndromes and these  
 CC pathological conditions are linked to Flavivirus infection or cancers.  
 CC The invention is also useful in gene therapy. The present sequence is a  
 CC p(95-114) EGFP (enhanced green fluorescent protein) (M1-M40) DEN (dengue)-2  
 CC (136F) plasmid DNA encoded protein.  
 XX  
 SQ Sequence 48 AA;  
 Query Match 91.1%; Score 51; DB 9; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 ETWFLRHP 9  
 DB 41 ETWFLRHP 48  
 |||||  
 RESULT 3  
 ADW12595  
 ID ADW12595 standard; peptide; 9 AA.  
 XX  
 AC ADW12595;  
 XX  
 DT 24-MAR-2005 (first entry)  
 XX  
 DE M32-40/DEN-2 mutant protein #1.  
 XX  
 KW Gene therapy; protein purification; virucide; cytostatic; vaccine;  
 KW hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;  
 KW DEN; dengue; mutant; mutein.





```

XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Catteau A;
XX XX WPI; 2005-047647/05.
XX PT New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
XX PT as a vaccine for preventing or treating pathological conditions or non-
XX PT specific febrile illnesses to severe hemorrhagic manifestations or
XX PT encephalitic syndromes.
XX PS Example 3; SEQ ID NO 25; 30pp; English.
XX CC The present invention relates to an isolated and purified ApoptoM
XX CC peptide. The invention is useful as a vaccine for the prevention and
XX CC treatment of pathological conditions from non-specific febrile illnesses
XX CC to severe hemorrhagic manifestations, encephalitic syndromes and these
XX CC pathological conditions are linked to Flavivirus infection or cancers.
XX CC The invention is also useful in gene therapy. The present sequence is a
XX CC M1-40/YF (yellow fever).17D (T34, I36, I37, H39) mutant protein.
XX SQ Sequence 40 AA;
Query Match 80.4%; Score 45; DB 9; Length 40;
Best Local Similarity 87.5%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ETWFLRHP 9
DB 33 ETWILRHP 40
|||||
RESULT 9
AAE17433
ID AAE17433 standard; protein; 48 AA.
XX AC AAE17433;
XX DT 18-APR-2002 (first entry)
XX DE (95-114)EGFP(206-245)DEN-2 fusion protein.
XX KW Dengue virus; pRM glycoprotein; E glycoprotein; apoptosis; virucide;
XX KW cancer; flavivirus infection; cytostatic; EGFP; DEN-2 protein;
XX KW enhanced green fluorescent protein; fusion protein; M ectodomain.
XX OS Dengue virus; 2.
XX OS Dengue virus; 1.
XX OS Unidentified.
XX OS Chimeric.
XX FH Key Location/Qualifiers
XX FT Misc-difference 13. .44 /note= "Encoded by GTATC"
XX FT
XX PN WO200196376-A2.
XX PD 20-DEC-2001.
XX PF 18-JUN-2001; 2001WO-IB001570.
XX PR 16-JUN-2000; 2000US-0212129P.
XX PA (INSP ) INST PASTEUR.
XX PI Despres P, Courageot M, Deubel V, Catteau A;
XX XX WPI; 2002-139706/18.
XX DR N-PSDB; AAD27335.
XX XX Novel apoptosis inducing polypeptide fragments of Dengue virus-1 or 2 M
XX PT protein, useful for inducing apoptosis in a cell of a human patient
PT suffering from cancer or flavivirus infection.
XX PS Claim 42; Fig 11; 45pp; English.
XX CC The invention relates to pro-apoptotic fragments of the Dengue virus
XX CC (DEN) pRM and E glycoproteins, methods for screening molecules capable of
XX CC inducing apoptosis and methods of inducing apoptosis in a cell. The
XX CC invention particularly relates to DEN-1 M (a membrane protein anchored in
XX CC envelope surrounding the nucleocapsid of the virus) ectodomain sequences,
XX CC Den-1-C amino acid sequence and DEN-2 M ectodomain sequence. Sequences of
XX CC the invention are useful for inducing apoptosis in a cell of a patient
XX CC suffering from cancer or flavivirus infection. They are also useful for
XX CC screening molecules which inhibit apoptosis. The present sequence is (95-
XX CC 114)EGFP(206-245)DEN-2 fusion protein construct. This construct comprises
XX CC 95-114 of the C-terminus of the C-protein of the DEN-1 virus strain BR/90
XX CC fused to the N-terminus of enhanced green fluorescent protein (EGFP) and
XX CC DEN-2 virus strain Jamaica M ectodomain (DEN-2 polypeptide) fused to the
XX CC C-terminus of the EGFP sequence
XX SQ Sequence 48 AA;
Query Match 80.4%; Score 45; DB 5; Length 48;
Best Local Similarity 87.5%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ETWFLRHP 9
DB 41 ETWILRHP 48
|||||
RESULT 10
ADN37497
ID ADN37497 standard; protein; 167 AA.
XX AC ADN37497;
XX DT 17-JUN-2004 (first entry)
XX DE Dengue virus C15/truncated pRM antigen fusion protein - SEQ ID 122.
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; truncated pRM;
XX KW capsid.
XX OS Dengue virus.
XX XX WO2003102166-A2.
XX PD 11-DEC-2003.
XX PF 26-FEB-2003; 2003WO-US005918.
XX PR 26-FEB-2002; 2002US-0360030P.
XX PA (MAXY-) MAXYGEN INC.
XX PI Apt D, Punnonen J, Brinkman AM;
XX XX WPI; 2004-043106/04.
XX PT New recombinant or synthetic polypeptides and polynucleotides useful for
XX PT diagnosing, preventing or treating diseases associated with flaviviruses,
XX PT including dengue viruses.
XX PS Disclosure; SEQ ID NO 122; 409pp; English.
XX CC The invention relates to a novel recombinant or synthetic polypeptide
XX CC comprising an amino acid sequence that has at least about 90% sequence
XX CC identity to any of the 20 fully defined amino acid sequences given in the
XX CC specification. The polypeptide of the invention demonstrates virucide
XX CC activity and may be useful for inducing an immune response to
XX CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
XX CC in detecting and/or diagnosing the presence of antibodies against the
XX CC Dengue virus serotypes in a sample and for gene therapy. The current

```

CC sequence is that of a Dengue virus C15/truncated prM antigen fusion  
 CC protein of the invention which comprises the C-terminal 15 amino acids of  
 CC the capsid protein fused to a truncated form of the prM protein lacking  
 CC the C-terminal 15 amino acids.

XX SQ Sequence 167 AA;

Query Match 80.4%; Score 45; DB 8; Length 167;  
 Best Local Similarity 87.5%; Pred. No. 5.1;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 ETWFLRHP 9  
 Db 140 ETWILRHP 147

RESULT 11

ADN37493  
 ID ADN37493 standard; protein; 171 AA.

XX AC ADN37493;

XX DT 17-JUN-2004 (first entry)

XX DE Dengue virus type 2 (DEN-2) C15/truncated prM antigen fusion protein.

XX KW virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;  
 XX KM capsid; DEN-2.

XX OS Dengue virus type 2.

XX PN W02003102166-A2.

XX PD 11-DEC-2003.

XX PF 26-FEB-2003; 2003WO-US005918.

XX PR 26-FEB-2002; 2002US-0360030P.

XX PA (MAXY-) MAXYGEN INC.

XX PI Apt D, Punnonen J, Brinkman AM;

XX DR WPI; 2004-043106/04.

XX PT New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

XX PS Disclosure; SEQ ID NO 118; 409pp; English.

XX CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of a Dengue virus type 2 (DEN-2) C15/truncated prM  
 CC antigen fusion protein of the invention which comprises the C-terminal 15  
 CC amino acids of the capsid protein fused to a truncated form of the prM  
 CC protein lacking the C-terminal 15 amino acids.

XX SQ Sequence 171 AA;

Query Match 80.4%; Score 45; DB 8; Length 171;  
 Best Local Similarity 87.5%; Pred. No. 5.2;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 ETWFLRHP 9  
 Db 144 ETWILRHP 151

RESULT 12

ADN37496  
 ID ADN37496 standard; protein; 171 AA.

XX AC ADN37496;

XX DT 17-JUN-2004 (first entry)

XX DE Dengue virus C15/truncated prM antigen fusion protein - SEQ ID 121.

XX KW virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;  
 XX KM capsid.

XX OS Dengue virus.

XX PN W02003102166-A2.

XX PD 11-DEC-2003.

XX PF 26-FEB-2003; 2003WO-US005918.

XX PR 26-FEB-2002; 2002US-0360030P.

XX PA (MAXY-) MAXYGEN INC.

XX PI Apt D, Punnonen J, Brinkman AM;

XX DR WPI; 2004-043106/04.

XX PT New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

XX PS Disclosure; SEQ ID NO 121; 409pp; English.

XX CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of a Dengue virus C15/truncated prM antigen fusion  
 CC protein of the invention which comprises the C-terminal 15 amino acids of  
 CC the capsid protein fused to a truncated form of the prM protein lacking  
 CC the C-terminal 15 amino acids.

XX SQ Sequence 171 AA;

Query Match 80.4%; Score 45; DB 8; Length 171;  
 Best Local Similarity 87.5%; Pred. No. 5.2;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 Db 144 ETWILRHP 151

RESULT 13

AAW75410  
 ID AAW75410 standard; peptide; 635 AA.

XX AC AAW75410;

XX DT 17-OCT-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 02-MAR-1999 (first entry)

XX DE Fusion protein PD30 contains Dengue virus epitope.



PF 26-FEB-2003; 2003WO-US005918.

PS Claim 40; SEQ ID NO 143; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.  
 XX  
 SQ Sequence 675 AA;

Query Match 80.4%; Score 45; DB 8; Length 675;  
 Best Local Similarity 87.5%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9  
 DB 140 ETWILRHP 147  
 ||| ||||

RESULT 16  
 ADN37612  
 ID ADN37612 standard; protein; 675 AA.  
 XX  
 AC ADN37612;  
 DT 17-JUN-2004 (first entry)  
 DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 237.  
 XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;  
 KW envelope; capsid.  
 XX Dengue virus.  
 OS WO2003102166-A2.  
 PN  
 PD 11-DEC-2003.  
 PF 26-FEB-2003; 2003WO-US005918.  
 PR 26-FEB-2002; 2002US-0360030P.  
 XX (MAXY-) MAXYGEN INC.  
 PA Apt D, Punnonen J, Brinkman AM;  
 XX WPI; 2004-043106/04.  
 DR  
 XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX Claim 40; SEQ ID NO 237; 409pp; English.  
 XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.  
 XX  
 SQ Sequence 675 AA;

Query Match 80.4%; Score 45; DB 8; Length 675;  
 Best Local Similarity 87.5%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9  
 DB 140 ETWILRHP 147  
 ||| ||||

RESULT 17  
 ADN37626  
 ID ADN37626 standard; protein; 675 AA.  
 XX  
 AC ADN37626;  
 DT 17-JUN-2004 (first entry)  
 DE Dengue virus C15/prM/E part codon-optimised antigen fusion protein 1.  
 XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E;  
 KW human codon-optimised; prM; envelope; capsid.  
 XX Dengue virus.  
 OS Synthetic.  
 PN WO2003102166-A2.  
 PD 11-DEC-2003.  
 PF 26-FEB-2003; 2003WO-US005918.  
 PR 26-FEB-2002; 2002US-0360030P.  
 XX (MAXY-) MAXYGEN INC.  
 PA Apt D, Punnonen J, Brinkman AM;  
 XX WPI; 2004-043106/04.  
 DR N-PSDB; ADN37630.  
 XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX Claim 40; SEQ ID NO 251; 409pp; English.  
 XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E partially human codon-  
 CC optimised antigen fusion protein of the invention which comprises 15  
 CC amino acids of the capsid (C) protein fused to the full-length partially  
 CC codon-optimised prM protein and envelope (E) protein.  
 XX  
 SQ Sequence 675 AA;

Query Match 80.4%; Score 45; DB 8; Length 675;  
 Best Local Similarity 87.5%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9  
 DB 140 ETWILRHP 147  
 ||| ||||

RESULT 18  
 AAW75411  
 ID AAW75411 standard; peptide; 677 AA.

```

XX AAW75411;
AC 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1999 (first entry)
XX
XX Fusion protein PD34 contains Dengue virus epitope.
DE
XX Dengue virus; fusion protein; P64K; Neisseria meningitidis; epitope;
KW antibody; diagnosis; Flavivirus; infection; vaccine.
XX
XX Dengue virus.
OS Neisseria meningitidis.
OS Chimeric.
XX
XX WO9831814-A1.
XX
XX 23-JUL-1998.
XX
XX 13-JAN-1998; 98WO-CU000001.
XX
XX 15-JAN-1997; 97CU-00000013.
XX
XX (CIGB-) CIGB CENT ING GENETICA & BIOTECNOLOGIA.
FA (IPKM-) IPK INST MEDICINA TROPICAL KOURI PEDRO.
XX
XX Vazquez Ramudo S, Guzman Tirado G, Guillen Nieto GE, Pardo Lazo OL;
PI Chinae Santiago G, Perez Diaz AB, Pupo Antunez M, Rodriguez Roche R;
PI Reyes Acosta O, Garay Perez HE, Padron Palomares G, Alvarez Vera M;
PI Morier Diaz L, Perez Insuaita O, Pelegrino Martinez De La Coterri Pedro;
XX
XX WPI; 1998-414111/35.
XX
XX New peptide(s) and fusion proteins useful for diagnosis and treatment of
PT flavivirus infection - contain cross-reactive epitopes from Dengue virus
PT pre-M/M protein and can induce neutralising antibodies.
XX
XX Claim 7; Page 30-32; 64pp; Spanish.
XX
XX This protein represents a fusion protein comprising an M protein epitope
CC from Dengue virus type 4 inserted into the P64K protein from Neisseria
CC meningitidis. Synthetic peptides based on the Dengue virus epitope
CC sequences (AAW75404-W75408) and fusion proteins can be used to raise
CC antibodies. The peptides, protein and antibodies are all useful for
CC diagnosis and treatment of Flavivirus infection, e.g. in vaccines.
CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 677 AA;
Query Match 80.4%; Score 45; DB 2; Length 677;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 117 ETWILRHP 124
|||||
117 ETWILRHP 124

RESULT 19
ADN37613
ID ADN37613 standard; protein; 677 AA.
XX
XX ADN37613;
AC
XX 17-JUN-2004 (first entry)
XX
XX Dengue virus C15/prM/E antigen fusion protein - SEQ ID 238.
DE
XX
XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;
KW envelope; capsid.
XX

```

```

OS Dengue virus.
XX WO2003102166-A2.
XX
XX 11-DEC-2003.
XX
XX 26-FEB-2003; 2003WO-US005918.
XX
XX 26-FEB-2002; 2002US-0360030P.
XX (MAXY-) MAXYGEN INC.
XX
XX Apt D, Punnonen J, Brinkman AM;
PI WPI; 2004-043106/04.
XX
XX New recombinant or synthetic polypeptides and polynucleotides useful for
PT diagnosing, preventing or treating diseases associated with flaviviruses,
PT including dengue viruses.
XX
XX Example 13; SEQ ID NO 238; 409pp; English.
XX
XX The invention relates to a novel recombinant or synthetic polypeptide
CC comprising an amino acid sequence that has at least about 90% sequence
CC identity to any of the 20 fully defined amino acid sequences given in the
CC specification. The polypeptide of the invention demonstrates virucide
CC activity and may be useful for inducing an immune response to
CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
CC in detecting and/or diagnosing the presence of antibodies against the
CC Dengue virus serotypes in a sample and for gene therapy. The current
CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of
CC the invention which comprises 15 amino acids of the capsid (C) protein
CC fused to the full-length prM protein and envelope (E) protein.
XX
XX Sequence 677 AA;
Query Match 80.4%; Score 45; DB 8; Length 677;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 140 ETWILRHP 147
|||||
140 ETWILRHP 147

RESULT 20
ADN37603
ID ADN37603 standard; protein; 681 AA.
XX
XX ADN37603;
AC
XX 17-JUN-2004 (first entry)
XX
XX Dengue virus type 2 Den-2C15/prM/E antigen fusion protein.
DE
XX
XX virucide; Flavivirus; arboviruses group B; gene therapy; DEN-2;
KW Den-2C15/prM/E; prM; envelope; capsid.
XX
XX Dengue virus type 2.
OS
XX WO2003102166-A2.
XX
XX 11-DEC-2003.
XX
XX 26-FEB-2003; 2003WO-US005918.
XX
XX 26-FEB-2002; 2002US-0360030P.
XX (MAXY-) MAXYGEN INC.
XX
XX Apt D, Punnonen J, Brinkman AM;
PI WPI; 2004-043106/04.
XX

```

```

XX New recombinant or synthetic polypeptides and polynucleotides useful for
PT diagnosing, preventing or treating diseases associated with flaviviruses,
PT including dengue viruses.
XX
PS Claim 38; SEQ ID NO 228; 409pp; English.
XX
CC The invention relates to a novel recombinant or synthetic polypeptide
CC comprising an amino acid sequence that has at least about 90% sequence
CC identity to any of the 20 fully defined amino acid sequences given in the
CC specification. The polypeptide of the invention demonstrates virucide
CC activity and may be useful for inducing an immune response to
CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
CC in detecting and/or diagnosing the presence of antibodies against the
CC Dengue virus serotypes in a sample and for gene therapy. The current
CC sequence is that of the Dengue virus type 2 (DEN-2) Den-2C15/prM/E
CC antigen fusion protein of the invention which comprises 15 amino acids of
CC the capsid (C) protein fused to the full-length prM protein and envelope
CC (E) protein.
XX
SQ Sequence 681 AA;

    Query Match      80.4%; Score 45; DB 8; Length 681;
    Best Local Similarity 87.5%; Pred. No. 24;
    Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 ETWFLRHP 9
Db      144 ETWILRHP 151
      ||| |||||

RESULT 21
ADN37517
ID ADN37517 standard; protein; 681 AA.
XX
AC ADN37517;
XX
DT 17-JUN-2004 (first entry)
XX
DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 142.
XX
KW virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;
KW envelope; capsid.
XX
OS Dengue virus.
XX
PN WO2003102166-A2.
XX
PD 11-DEC-2003.
XX
PF 26-FEB-2003; 2003WO-US005918.
XX
PR 26-FEB-2002; 2002US-0360030P.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Apt D, Punnonen J, Brinkman AM;
XX
DR WPI; 2004-043106/04.
XX
PT New recombinant or synthetic polypeptides and polynucleotides useful for
PT diagnosing, preventing or treating diseases associated with flaviviruses,
PT including dengue viruses.
XX
PS Claim 40; SEQ ID NO 142; 409pp; English.
XX
CC The invention relates to a novel recombinant or synthetic polypeptide
CC comprising an amino acid sequence that has at least about 90% sequence
CC identity to any of the 20 fully defined amino acid sequences given in the
CC specification. The polypeptide of the invention demonstrates virucide
CC activity and may be useful for inducing an immune response to
CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
CC in detecting and/or diagnosing the presence of antibodies against the

```

---

```

CC Dengue virus serotypes in a sample and for gene therapy. The current
CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of
CC the invention which comprises 15 amino acids of the capsid (C) protein
CC fused to the full-length prM protein and envelope (E) protein.
XX
SQ Sequence 681 AA;

    Query Match      80.4%; Score 45; DB 8; Length 681;
    Best Local Similarity 87.5%; Pred. No. 24;
    Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 ETWFLRHP 9
Db      144 ETWILRHP 151
      ||| |||||

RESULT 22
ABP57874
ID ABP57874 standard; protein; 685 AA.
XX
AC ABP57874;
XX
DT 07-FEB-2003 (first entry)
XX
DE Plasmid pCBD2-14-6 containing dengue-2 virus prM and E.
XX
KW Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;
KW pCBD2-14-6; dengue virus; DEN-2.
XX
OS Unidentified.
OS Dengue-2 virus.
OS Chimeric.
XX
PN WO200281754-A1.
XX
PD 17-OCT-2002.
XX
PF 04-APR-2002; 2002WO-US010764.
XX
PR 04-APR-2001; 2001US-00826115.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Chang GJ;
XX
DR WPI; 2003-058572/05.
DR N-PSDB; ABV77547.
XX
PT Novel isolated nucleic acid useful as vaccine for preventing flavivirus
PT infection, comprises transcriptional unit encoding signal sequence of one
PT flavivirus and immunogenic flavivirus antigen of a second flavivirus.
XX
PS Example 20; Page 157-158; 174pp; English.
XX
CC The invention relates to a novel nucleic acid comprising a
CC transcriptional unit encoding a signal sequence of a structural protein
CC of a first flavivirus and an immunogenic flavivirus antigen of a second
CC flavivirus, where the transcriptional unit directs the synthesis of the
CC antigen. The polynucleotide of the invention has virucide activity, and
CC acts as a vaccine. A composition of the invention is useful for
CC immunising a subject against infection by a flavivirus. The
CC polynucleotide is useful as a vaccine for preventing flavivirus
CC infection. The sequence represents plasmid pCBD2-14-6, which contains
CC dengue-2 virus (DEN-2) prM and E proteins
XX
SQ Sequence 685 AA;

    Query Match      80.4%; Score 45; DB 6; Length 685;
    Best Local Similarity 87.5%; Pred. No. 24;
    Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 ETWFLRHP 9
      ||| |||||

```

```

Db      148 ETWILRHP 155

RESULT 23
ABP57876
ID      ABP57876 standard; protein; 685 AA.
XX
AC      ABP57876;
XX
DT      07-FEB-2003 (first entry)
XX
DE      Plasmid pCB8D2-2J-2-9-1 protein product.
XX
KW      Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;
KW      pCB8D2-2J-2-9-1; Japanese encephalitis virus; dengue-2 virus; DEN-2.
XX
OS      Unidentified.
OS      Synthetic.
XX
PN      WO200281754-A1.
XX
PD      17-OCT-2002.
XX
PF      04-APR-2002; 2002WO-US010764.
XX
PR      04-APR-2001; 2001US-00826115.
XX
PA      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI      Chang GJ;
XX
PX      WPI; 2003-058572/05.
XX
DR      N-PSDB; ABV77548.
XX
PT      Novel isolated nucleic acid useful as vaccine for preventing flavivirus
PT      infection, comprises transcriptional unit encoding signal sequence of one
PT      flavivirus and immunogenic flavivirus antigen of a second flavivirus.
XX
PS      Example 20; Page 162-164; 174pp; English.
XX
CC      The invention relates to a novel nucleic acid comprising a
CC      transcriptional unit encoding a signal sequence of a structural protein
CC      of a first flavivirus and an immunogenic flavivirus antigen of a second
CC      flavivirus, where the transcriptional unit directs the synthesis of the
CC      antigen. The polynucleotide of the invention has virucide activity, and
CC      acts as a vaccine. A composition of the invention is useful for
CC      immunising a subject against infection by a flavivirus. The
CC      polynucleotide is useful as a vaccine for preventing flavivirus
CC      infection. The sequence represents plasmid pCB8D2-2J-2-9-1, which
CC      contains dengue-2 virus (DEN-2) prM, M and E, and Japanese encephalitis
CC      virus E proteins
XX
SQ      Sequence 685 AA;

Query Match      80.4%; Score 45; DB 6; Length 685;
Best Local Similarity 87.5%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 ETWFLRHP 9
        ||| ||||
Db      148 ETWILRHP 155

RESULT 25
AAW09409
ID      AAW09409 standard; protein; 1127 AA.
XX
AC      AAW09409;
XX
DT      17-OCT-2003 (revised)
DT      19-MAY-1997 (first entry)
XX
DE      Dengue virus serotype 2 PR159/S1 polypeptide.
XX
KW      DEN-2; flavivirus; envelope protein; immunisation; vaccine.
XX
OS      Dengue virus; serotype 2.
XX
Key      Location/Qualifiers
FH      Region 1..114
FT      /label= Capsid
FT      Region 115..205
FT      /label= Pre-membrane
FT      Region 206..280
FT      /label= Membrane
FT      Region 281..775

```

```

FT      Domain      /label= Envelope
FT      296..395
FT      /label= Domain-B
FT      Misc-difference 588
FT      /note= "amino acid residue 588 (Val) is Ile in wild-type
FT      PR159"
FT      Region      776..1127
FT      /label= NS1
XX
XX      WO9637221-A1.
XX
XX      28-NOV-1996.
XX
XX      24-MAY-1996; 96WO-US007627.
XX
XX      24-MAY-1995; 95US-00448734.
XX      07-JUN-1995; 95US-00488807.
XX      10-JUL-1995; 95US-00500469.
XX
XX      (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.
XX
XX      Ivy JM, Nakano E, Clements D;
XX      WPI; 1997-020938/02.
XX      N-PSDB; AAT47666.
XX
XX      Sub:unit vaccine against flavivirus infection - contg. recombinant
XX      envelope protein in secretable form, used for immunising against
XX      flavivirus infection.
XX
XX      Example 1; Fig 3A-D; 121pp; English.
XX
XX      A polypeptide (AAW09409) comprises the capsid, pre-membrane, envelope and
XX      NS1 proteins of dengue virus serotype 2 (DEN-2) variant PR159/S1. A
XX      conservative mutation in the envelope protein may be involved in the
XX      attenuation of this small-plaque, temp.- sensitive variant. Portions of
XX      the envelope protein, esp. domain B, can be expressed in eukaryotic hosts
XX      (see also AAW09410 and AAW09427-28) transfectected with vectors
XX      incorporating DEN-2 S1 cDNA (see also AAT47666). These polypeptides can
XX      be used in novel subunit vaccines against viral infection, to raise
XX      antibodies useful for passive immunisation, and for diagnosis of
XX      infection. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX      Sequence 1127 AA;
XX
XX      Query Match      80.4%; Score 45; DB 2; Length 1127;
XX      Best Local Similarity 87.5%; Pred. No. 41;
XX      Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
Qy      2 ETWFLRHP 9
Db      238 ETWILRHP 245

RESULT 26
ID      AAY05522 standard; protein; 1127 AA.
XX
XX      AAY05522;
XX
XX      17-OCT-2003 (revised)
XX      05-JUL-1999 (first entry)
XX
XX      Dengue virus serotype 2 PR159/S1 viral capsid, pprM, E, NS1.
XX      Flavivirus; envelope protein; vaccine; infection; diagnosis.
XX
XX      Dengue virus; serotype 2.
XX
XX      Key      Location/Qualifiers
XX      Protein 1..114
XX      /label= Capsid
XX      115..205
XX
XX      Dengue virus type 2; strain PR159/S1.

```

---

```

FT      Protein      /label= PreMembrane
FT      206..280
FT      /label= Membrane
FT      280..1127
FT      /label= Envelope
XX
XX      WO9906068-A2.
XX
XX      11-FEB-1999.
XX
XX      27-JUL-1998; 98WO-US015447.
XX
XX      31-JUL-1997; 97US-00904227.
XX
XX      (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.
XX
XX      Ivy JM, Peters ID, Collier BG, McDonnell M, Harada KE;
XX      WPI; 1999-153454/13.
XX      N-PSDB; AAX25114.
XX
XX      Recombinant dimeric flaviviral envelope vaccine - comprising a dimeric
XX      80%E protein, useful for protecting against flavivirus, especially dengue
XX      virus infections.
XX
XX      Example 1; Fig 3A-D; 60pp; English.
XX
XX      This sequence is composed of the capsid, prM, envelope (E) and NS1
XX      proteins of serotype 2 dengue virus DEN-2 strain PR159/S1. A vaccine for
XX      protecting against flavivirus infection comprises a dimeric 80% E protein
XX      that has been secreted as a recombinant protein from a eukaryotic cell.
XX      80% E indicates a C-terminally truncated flavivirus E protein. The
XX      dimeric truncated E is formed: (1) by directly linking 2 tandem copies of
XX      80% E via a flexible tether; (2) via the formation of a leucine zipper
XX      domain through the homodimeric association of 2 leucine zipper helices
XX      each fused to the C-terminus of an 80% E molecule; or (3) via the
XX      formation of a non-covalently associated four-helix bundle domain formed
XX      upon association of two helix-turn-helix moieties attached to the C-
XX      terminus of an 80% E molecule. Dimeric truncated DEN-2 E proteins are
XX      efficiently secreted by recombinant cells, are easier to purify than
XX      intracellular proteins, and generate a high titer neutralising antibody
XX      response. The method is generally applicable to flaviviruses, in
XX      particular dengue viruses such as DEN-2, where 80% E comprises amino
XX      acids 1-395 of DEN-2 E. The products can also be used for diagnosis of
XX      infection. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX      Sequence 1127 AA;
XX
XX      Query Match      80.4%; Score 45; DB 2; Length 1127;
XX      Best Local Similarity 87.5%; Pred. No. 41;
XX      Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
Qy      2 ETWFLRHP 9
Db      238 ETWILRHP 245

RESULT 27
ID      ADL98086
XX
XX      ADL98086 standard; protein; 1127 AA.
XX
XX      ADL98086;
XX
XX      18-NOV-2004 (first entry)
XX
XX      Dengue virus, DEN-2, capsid/membrane/envelope/NS1 proteins.
XX
XX      Dengue virus; DEN-2; Envelope protein; 80% E; membrane protein;
XX      capsid protein; NS1 protein; Dengue haemorrhagic fever; DHF;
XX      Dengue shock syndrome; DSS; flavivirus; vaccine.
XX
XX      Dengue virus type 2; strain PR159/S1.

```

PN US2003175304-A1.  
 XX 18-SEP-2003.  
 XX  
 XX 20-SEP-2002; 2002US-00247960.  
 XX  
 XX 31-JUL-1997; 97US-00904227.  
 PR 18-AUG-1999; 99US-00376463.  
 XX  
 XX (PETE/) PETERS I D.  
 PA (COLL/) COLLIER B G.  
 PA (MCDON/) MCDONELL M.  
 PA (IVYJ/) IVY J M.  
 PA (HARA/) HARADA K.  
 XX  
 XX Peters ID, Collier BG, McDonnell M, Ivy JM, Harada K;  
 XX  
 XX WPI; 2003-898503/82.  
 DR N-PSDB; ADL98085.  
 XX  
 XX Vaccine useful for protection against dengue virus infection, comprises a  
 PT dimeric 80% envelope, which has been secreted as a recombinantly produced  
 PT protein from Drosophila Schneider cells.  
 XX  
 XX Example 1; Fig 3; 31pp; English.  
 PS  
 XX The invention relates to a vaccine for protection against Flavivirus  
 XX infection comprising a dimeric 80% envelope (E), which has been secreted  
 CC as a recombinantly produced protein from Drosophila Schneider cells and  
 CC which represents the N-terminal 80% portion of the protein from residue 1  
 CC -395. Also included are a method for protecting a subject against a  
 CC Flavivirus, an immunogenic polypeptide comprising a dimeric 80% E, an  
 CC immunogenic composition for protection against Flavivirus infection  
 CC comprising the immunogenic polypeptide and a carrier, an immunodiagnostic  
 CC for detecting Flavivirus comprising the immunogenic polypeptide, a vector  
 CC host recombinant DNA expression system, a DNA sequence encoding the  
 CC immunogenic polypeptide and an immunodiagnostic kit for detecting  
 CC Flavivirus in a test subject. The dimeric 80% E products are envelope  
 CC proteins of serotypes comprising DEN-1, DEN-2, DEN-3 or DEN-4. The  
 CC Flavivirus is a dengue virus. The 80% E protein is produced as a dimer by  
 CC incorporating 2 different kinds of leucine zipper peptides or  
 CC incorporating a helix-turn-helix peptide, to encourage dimerisation. The  
 CC vaccine is useful for protection against dengue virus infection (e.g.  
 CC dengue haemorrhagic fever, DHF, and dengue shock syndrome, DSS). The  
 CC present sequence is encoded by the partial genomic sequence of the DEN-2  
 CC strain PR159/S1 virus, and represents the capsid, membrane, envelope and  
 CC NS1 proteins.  
 XX  
 XX Sequence 1127 AA;  
 SQ  
 Query Match 80.4%; Score 45; DB 7; Length 1127;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 ETWFLRHP 9  
 Db 238 ETWILRHP 245  
 RESULT 28  
 ADQ28716  
 ID ADQ28716 standard; protein; 1127 AA.  
 XX  
 XX AC ADQ28716;  
 XX  
 XX 26-AUG-2004 (first entry)  
 DT  
 XX Dengue virus viral capsid, prM, E and NS1 gene polyprotein.  
 DE  
 XX virucide; vaccine; Flavivirus; dimeric 80%; Drosophila Schneider cell;  
 KW immunogenic composition; multivalent immunodiagnostic; dengue virus;  
 KW viral capsid; prM gene; E gene; NS1 gene.  
 XX

OS Dengue virus.  
 XX US6749857-B1.  
 PN  
 XX 15-JUN-2004.  
 PD  
 XX 18-AUG-1999; 99US-00376463.  
 XX  
 XX 31-JUL-1997; 97US-00904227.  
 PR  
 XX (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.  
 PA  
 XX Peters ID, Collier BG, McDonnell M, Ivy JM, Harada K;  
 PI  
 XX WPI; 2004-438725/41.  
 DR N-PSDB; ADQ28715.  
 DR  
 XX New vaccines for preventing or diagnosing infections caused by dengue  
 PT virus comprises a therapeutic amount of a dimeric 80% protein secreted  
 PT from Drosophila Schneider cells.  
 XX  
 XX Example 1; SEQ ID NO 3; 47pp; English.  
 PS  
 XX The invention describes a vaccine that generates a protective,  
 CC neutralising antibody response to a Flavivirus in a murine host. The  
 CC vaccine comprises a therapeutic amount of a dimeric 80% E, the dimeric  
 CC 80% E having been secreted as a recombinantly produced protein from  
 CC Drosophila Schneider cells, and where 80% E represents the N-terminal 80%  
 CC portion of the protein from residues 1-395. Also described are: an  
 CC immunogenic polypeptide comprising the dimeric 80% E cited above; an  
 CC immunogenic composition that generates a protective, neutralising  
 CC antibody response to a Flavivirus in a murine host, comprising the above  
 CC immunogenic polypeptide and a physiological carrier; a multivalent  
 CC immunodiagnostic for the detection of Flavivirus, comprising at least 2  
 CC of the above immunogenic polypeptides of at least 2 flavivirus serotypes;  
 CC and an immunodiagnostic kit for the detection of Flavivirus in a test  
 CC subject, comprising the above immunogenic or multivalent immunodiagnostic  
 CC polypeptide, a suitable support phase coated with dimeric 80% E, and  
 CC labeled antibodies immunoreactive to antibodies from the test subject.  
 CC The composition is useful for preventing or diagnosing infections caused  
 CC by dengue virus. This is the amino acid sequence of the polyprotein  
 CC encoded by dengue virus gene viral capsid, prM, E and NS1 genes for  
 CC Dengue virus strain PR159/S1 used as the source of DEN-2 genes for the  
 CC invention.  
 XX  
 XX Sequence 1127 AA;  
 SQ  
 Query Match 80.4%; Score 45; DB 8; Length 1127;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 ETWFLRHP 9  
 Db 238 ETWILRHP 245  
 RESULT 29  
 AAE35314  
 ID AAE35314 standard; protein; 3388 AA.  
 XX  
 XX AC AAE35314;  
 XX  
 XX 28-MAY-2003 (first entry)  
 DT  
 XX Dengue virus type 2 strain rDEN2/4delta30 protein.  
 DE  
 XX Attenuation; growth; vaccine; infection; Dengue virus type 4.  
 XX  
 XX Dengue virus.  
 OS  
 XX WO200295075-A1.  
 PN  
 XX 28-NOV-2002.  
 PD

```

XX 22-MAY-2002; 2002MO-US016308.
PF XX
PR XX
XX 22-MAY-2001; 2001US-0293049P.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (BLAN/) BLANEY J E.
XX
XX Whitehead SS, Murphy BR, Hanley KA;
XX WPI; 2003-120809/11.
XX N-PSDB; AAD53912.
XX
XX New mutated flavivirus, useful for fine tuning the attenuation and growth
XX characteristics of dengue virus vaccines for the prevention and/or
XX treatment of dengue virus infection.
XX
XX Disclosure; Page 133-134; 246pp; English.
XX
XX The present invention relates to novel mutated flaviviruses comprising a
XX phenotype in which the viral genome is modified by introduction of a
XX mutation, singly or in combination, taken from mutations from recombinant
XX virus bearing Vero adaptation mutations, putative Vero cell adaptation
XX mutations of dengue type 4 virus (DENV4) or mutations known to attenuate
XX dengue type 4 virus. The methods and compositions of the invention are
XX useful for fine tuning the attenuation and growth characteristics of
XX dengue virus vaccines for the prevention and/or treatment of dengue virus
XX infection. The present sequence is Dengue virus type 4 strain
XX rDENV2/4delta30 protein
XX
XX Sequence 3388 AA;
XX
XX Query Match 80.4%; Score 45; DB 6; Length 3388;
XX Best Local Similarity 87.5%; Pred. No. 1.3e+02;
XX Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2 ETWFLRHP 9
DB 238 ETWILRHP 245
XX
RESULT 30
AAR13166
ID AAR13166 standard; protein; 3391 AA.
XX
AC AAR13166;
XX
XX 25-MAR-2003 (revised)
DT 21-NOV-1991 (first entry)
XX
XX Proteins encoded by entire Dengue 2 virus genome.
XX
XX dengue virus; detection; consensus sequence; Flavivirus; PCR.
XX
XX Dengue virus.
XX
XX Key Location/Qualifiers
XX Peptide 116..205
XX /label= prM
XX Modified-site 183
XX /label= N-glycosylated
XX Protein 206..280
XX /label= M
XX Protein 281..775
XX /label= E
XX Modified-site 347
XX /label= N-glycosylated
XX Modified-site 433
XX /label= N-glycosylated
XX Protein 776..1127
XX /label= NS1
XX Modified-site 905
XX /label= N-glycosylated

```

```

FT Modified-site 982
FT /label= N-glycosylated
FT Protein 1128..1345
FT /label= NS2A
FT Modified-site 1134
FT /label= N-glycosylated
FT Modified-site 1174
FT /label= N-glycosylated
FT Modified-site 1329
FT /label= N-glycosylated
FT Protein 1346..1474
FT /label= NS2B
FT Modified-site 1369
FT /label= N-glycosylated
FT Protein 1475..2093
FT /label= NS3
FT Protein 2094..2243
FT /label= ns4a
FT Protein 2244..2492
FT /label= NS4B
FT Modified-site 2301
FT /label= N-glycosylated
FT Modified-site 2305
FT /label= N-glycosylated
FT Modified-site 2457
FT /label= N-glycosylated
FT Modified-site 2485
FT /label= N-glycosylated
FT Protein 2493..3391
FT /label= NS5
FT Modified-site 2644
FT /label= N-glycosylated
FT Modified-site 2865
FT /label= N-glycosylated
FT Modified-site 2704
FT /label= N-glycosylated
FT Modified-site 2714
FT /label= N-glycosylated
XX
XX FR2654113-A.
XX
XX 10-MAY-1991.
XX
XX 09-NOV-1989; 89FR-00914724.
XX
XX 09-NOV-1989; 89FR-00014724.
XX
XX (INSP ) INST PASTEUR.
XX
XX Vincent D;
XX
XX WPI; 1991-225002/31.
XX N-PSDB; AAQ12787.
XX
XX Detection and identification of Flaviviridae in biological sample - by
XX amplifying consensus sequence then hybridisation opt. followed by typing,
XX e.g. sequencing amplified prod.
XX
XX Disclosure; Fig 3; 24pp; French.
XX
XX The dengue 2 virus is an example of a member of the Flaviviridae which
XX can be identified using the probe pair of the invention. A species-
XX specific sequence can be amplified using the claimed oligonucleotides as
XX primers in a PCR reaction (see AAQ12788 and AAQ12789). Other viruses
XX which can be identified include Japanese encephalitis virus and yellow
XX fever virus. All the dengue 2 virus proteins are encoded from an
XX uninterrupted genomic sequence. (Updated on 25-MAR-2003 to correct PR
XX field.)
XX
XX Sequence 3391 AA;
SQ

```

```

Query Match 80.4%; Score 45; DB 2; Length 3391;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;

```



---

Matches	7;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	2	ETWFLRHP	9						
Db	238	ETWILRHP	245						

Search completed: August 31, 2006, 11:50:34  
Job time : 112.25 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 11:43:31 ; Search time 17.25 Seconds  
(without alignments)  
50.200 Million cell updates

```
Title:      DENGUE_SEROTYPE3
Perfect score: 56
Sequence:   1 retwflrhp 9
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10
                  Listing first 10

```

```
Database :      PIR_80:*
1:  _pir1:
2:  _pir2:
3:  _pir3:
4:  _pir4:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	80.4	166	2	S40144	premembrane protein
2	45	80.4	555	2	JQ1404	genome polyprotein
3	45	80.4	775	2	A48644	polyprotein - deng
4	45	80.4	1127	1	GNWVD2	genome polyprotein
5	45	80.4	3388	1	GNWVDP	genome polyprotein
6	45	80.4	3391	1	GNWV16	genome polyprotein
7	45	80.4	3391	1	GNWV26	genome polyprotein
8	45	80.4	3391	1	GNWVJA	genome polyprotein
9	45	80.4	3391	2	JS0219	polyprotein - deng
10	43	76.8	555	2	JQ1405	genome polyprotein
11	43	76.8	775	2	A47311	polyprotein(C, E,
12	43	76.8	792	2	C32401	genome polyprotein
13	43	76.8	792	2	B32401	genome polyprotein
14	43	76.8	792	2	A32401	genome polyprotein
15	43	76.8	1226	1	GNWVWP	genome polyprotein
16	43	76.8	3390	1	GNWVD3	genome polyprotein
17	43	76.8	3396	1	A42551	genome polyprotein
18	42	75.0	166	2	S03223	membrane protein -
19	42	75.0	166	2	S03225	membrane protein -
20	41	73.2	422	2	A83184	probable protein m
21	40	71.4	166	2	S03224	membrane protein -
22	39	69.6	205	2	E86085	hypothetical prote
23	39	69.6	205	2	A98238	hypothetical prote
24	39	69.6	343	2	H95879	probable sugar ABC
25	39	69.6	665	2	PS0043	genome polyprotein
26	38	67.9	205	2	I78665	hypothetical 23.0K
27	38	67.9	490	2	I41293	EcoE type I restri
28	38	67.9	1155	2	B96761	probable protein k
29	37	66.1	1984	1	S49184	phosphinothricin N

## ALIGNMENTS

```
RESULT 1
S40144
genome polypeptide - dengue virus type 2
C;Species: dengue virus type 2
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 31-Dec-2004
C;Accession: S40144
R;Shiu, S.Y.W.
submitted to the EMBL Data Library, May 1993
A;Reference number: S40144
A;Accession: S40144
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-166 <SHI>
A;Cross-references: UNIPROT:Q66346; UNIPARC:UPI00000F6DD9; EMBL:X72849; NID:G437772; PID:
C;Superfamily: hepatitis C virus genome polypeptide

Query Match      80.4%; Score 45; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 ETWFLRHP 9
Db      124 ETWILRHP 131

RESULT 2
JQ1404
genome polypeptide - dengue virus type 2 (strain TH-36) (fragment)
N;Contains: envelope protein E; membrane-associated protein M; nonstructural protein NS1
C;Species: dengue virus type 2
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004
C;Accession: JQ1404
R;Shiu, S.Y.W.; Jiang, W.R.; Porterfield, J.S.; Gould, E.A.
J. Gen. Virol. 73, 207-212, 1992
A;Title: Envelope protein sequences of dengue virus isolates TH-36 and TH-Sman, and identical
A;Reference number: JQ1404; MUID:92113574; PMID:11339466
A;Accession: JQ1404
A;Molecule type: genomic RNA
A;Residues: 1-555 <SHI>
A;Cross-references: UNIPROT:P29984; UNIPARC:UPI0000131DF8; GB:D10514; DDBJ:D01074; NID:
C;Superfamily: hepatitis C virus genome polypeptide
C;Keywords: envelope protein; glycoprotein; nonstructural protein; polypeptide; transmembrane
F;1-49/Product: membrane-associated protein M (fragment) #status predicted <MEM>
F;37-53/Domain: transmembrane #status predicted <TM1>
F;50-544/Product: envelope protein E #status predicted <ENV>
F;496-512/Domain: transmembrane #status predicted <TM2>
F;526-542/Domain: transmembrane #status predicted <TM3>
F;545-555/Product: nonstructural protein NS1 (fragment) #status predicted <NON>
F;116,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      80.4%; Score 45; DB 2; Length 555;
Best Local Similarity 87.5%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 ETWFLRHP 9
Db      7 ETWILRHP 14

RESULT 3
A48644
polypeptide - dengue virus type 2 (strain Mexican) (fragment)
C;Species: dengue virus type 2
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 31-Dec-2004
C;Accession: A48644
R;Ruiz, B.H.; Sanchez, I.; Ortega, G.J.; Lopez, I.; Ortiz-Ortiz, L.
submitted to GenBank, October 1992
A;Description: Nucleotide sequence and deduced amino-acid sequence of the structural protein
A;Reference number: A48644
A;Accession: A48644
```

```
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-775 <RUI>
A;Cross-references: UNIPROT:Q66398; UNIPARC:UPI00000EEB45; GB:L04561; NID:G323652; PIDN:
C;Superfamily: hepatitis C virus genome polypeptide
C;Keywords: polypeptide

Query Match      80.4%; Score 45; DB 2; Length 775;
Best Local Similarity 87.5%; Pred. No. 6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 ETWFLRHP 9
Db      238 ETWILRHP 245

RESULT 4
GNWVD2
genome polypeptide - dengue virus type 2 (strain D2-04) (fragment)
N;Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstructural
C;Species: dengue virus type 2
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 31-Dec-2004
C;Accession: JCI1007; JCI1005
R;Yang, P.Y.; Lam, S.K.
Chinese J. Microbiol. Immunol. 11, 341-344, 1991
A;Title: The nucleotide and encoded amino acid sequences of the structural protein gene
A;Reference number: JCI1007
A;Accession: JCI1007
A;Molecule type: genomic RNA
A;Residues: 1-775 <YAN>
A;Cross-references: UNIPROT:P30026; UNIPARC:UPI0000174A06
A;Note: the authors translated the codons TTA for residue 53 as Phe, AGT for residue 136
S as Arg, GGC for residue 266 as Ala, and CAG for residue 272 as Leu
R;Fan, P.Y.; Kautner, I.M.; Koh, C.L.; Lam, S.K.
Chinese J. Microbiol. Immunol. 11, 9-12, 1991
A;Title: Nucleotide and encoded amino acid sequences of the nonstructural protein NS1 gene
A;Reference number: JCI1005
A;Accession: JCI1005
A;Molecule type: genomic RNA
A;Residues: 776-1127 <YAZ>
A;Cross-references: UNIPARC:UPI0000174A07
A;Note: the authors translated the codons GTG for residue 899 as Leu, CTG for residue 95;
C;Superfamily: hepatitis C virus genome polypeptide
C;Keywords: capsid protein; envelope protein; glycoprotein; membrane-associated protein;
F;114/Product: capsid protein C #status predicted <CAP>
F;101-117/Domain: transmembrane #status predicted <TM1>
F;101-117/Domain: transmembrane-associated protein M precursor #status predicted <MAM>
F;115-205/Domain: nonterminal signal sequence #status predicted <SIG>
F;206-280/Product: membrane-associated protein M #status predicted <MEM>
F;281-775/Product: envelope protein E #status predicted <ENV>
F;727-743/Domain: transmembrane #status predicted <TM2>
F;757-773/Domain: transmembrane #status predicted <TM3>
F;776-1127/Product: nonstructural protein NS1 #status predicted <NPN>
F;183,347,433,905,982/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      80.4%; Score 45; DB 1; Length 1127;
Best Local Similarity 87.5%; Pred. No. 8.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 ETWFLRHP 9
Db      238 ETWILRHP 245

RESULT 5
GNWVDP
genome polypeptide - dengue virus type 2 (strain PR159/S1)
N;Contains: capsid protein; envelope protein; membrane protein; nonstructural protein NS5
a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: dengue virus type 2
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 31-Dec-2004
C;Accession: A29972
R;Hahn, Y.S.; Galler, R.; Hunkapiller, T.; Dalrymple, J.M.; Strauss, J.H.; Strauss, E.G.
```

```

Query Match      80.4%; Score 45; DB 1; Length 3391;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 ETWFLRHP 9
      ||| ||||
Db      238 ETWILRHP 245

RESULT 7
GNMV26
genome polyprotein - dengue virus type 2 (strain 16681-PDK53)
N;Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru-
tural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: dengue virus type 2
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 31-Dec-2004
R;Accession: B42451
R;Blok, J.; McWilliam, S.M.; Butler, H.C.; Gibbs, A.J.; Weiller, G.; Herring, B.L.; Hems-
Virology 187, 573-590, 1992
A;Title: Comparison of a dengue-2 virus and its candidate vaccine derivative: sequence r
A;Reference number: A42451; MUID:92189532; PMID:1312269
A;Accession: B42451
A;Molecule type: Genomic RNA
A;Residues: 1-3391 <BLO>
A;Cross-references: UNIPROT:P29991; UNIPARC:UPI000131DF6; GB:M85259
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
F/1-114/Product: capsid protein C #status predicted <CPC>
F/50-66/Domain: transmembrane #status predicted <TM1>
F/102-118/Domain: transmembrane #status predicted <TM2>
F/115-280/Product: membrane-associated protein M precursor #status predicted <MPP>
F/115-205/Domain: nonterminal signal sequence #status predicted <SIG>
F/206-280/Product: membrane-associated protein M #status predicted <MPM>
F/268-284/Domain: transmembrane #status predicted <TM3>
F/281-775/Product: envelope protein E #status predicted <EPE>
F/727-743/Domain: transmembrane #status predicted <TM4>
F/757-773/Domain: transmembrane #status predicted <TM5>
F/776-1127/Product: nonstructural protein NS1 #status predicted <NS1>
F/1128-1345/Product: nonstructural protein NS2a #status predicted <N2a>
F/1158-1174/Domain: transmembrane #status predicted <TM6>
F/1272-1288/Domain: transmembrane #status predicted <TM7>
F/1294-1310/Domain: transmembrane #status predicted <TM8>
F/1346-1474/Product: nonstructural protein NS2b #status predicted <N2B>
F/1351-1367/Domain: transmembrane #status predicted <TM9>
F/1373-1389/Domain: transmembrane #status predicted <TMa>
F/1448-1464/Domain: transmembrane #status predicted <TMb>
F/1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>

```

```

F:1759--1762/Region: DEAR_MocII1
F:2094-2243/Product: nonstructural protein NS4a #status predicted <N4A>
F:2148-2164/Domain: transmembrane #status predicted <TMC>
F:2174-2190/Domain: transmembrane #status predicted <TMD>
F:2197-2213/Domain: transmembrane #status predicted <TME>
F:2227-2243/Domain: transmembrane #status predicted <TMF>
F:2244-2491/Product: nonstructural protein NS4b #status predicted <N4B>
F:2352-2368/Domain: transmembrane #status predicted <TMG>
F:2411-2427/Domain: transmembrane #status predicted <TMH>
F:2492-3391/Product: nonstructural protein NS5 #status predicted <NS5>
F:183,347,433,905,982,1134,1174,1329,2301,2305,2346,2387,2457,2485,2644,2665,2704,2714/B

```



A47311  
 Polyprotein(C, E, M, prM) - dengue virus type 1 (fragment)  
 C/Species: dengue virus type 1  
 C/Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 31-Dec-2004  
 C/Accession: A47311  
 R/Despres, P.; Frenkiel, M.P.; Deubel, V.  
 Virology 196, 209-219, 1993  
 A/Title: Differences between cell membrane fusion activities of two dengue type-1 isolat  
 A/Reference number: A47311; MUID:93362407; PMID:8356794  
 A/Contents: BR/90  
 A/Accession: A47311  
 A/Status: preliminary  
 A/Molecule type: nucleic acid  
 A/Residues: 1-775 <DS>  
 A/Cross-references: UNIPROT:Q86647; UNIPARC:UPI00000EDGDA; GB:S64849; NID:G408338; PIDN:  
 A/Note: sequence extracted from NCBI backbone (NCBIN:136589, NCBIP:136590)  
 C/Superfamily: hepatitis C virus genome polyprotein  
 C/Keywords: polyprotein

Query Match 76.8%; Score 43; DB 2; Length 775;  
 Best Local Similarity 87.5%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 ||| |||||  
 Db 238 ETWALRHP 245

RESULT 12  
 C32401  
 Genome polyprotein - dengue virus type 1 (strain 836-1) (fragment)  
 N/Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru  
 C/Species: dengue virus type 1  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
 C/Accession: C32401  
 R/Chu, M.C.; O'Rourke, E.J.; Trent, D.W.  
 J. Gen. Virol. 70, 1701-1712, 1989  
 A/Title: Genetic relatedness among structural protein genes of dengue 1 virus strains.  
 A/Reference number: A32401; MUID:89293078; PMID:2738579  
 A/Accession: C32401  
 A/Molecule type: genomic RNA  
 A/Residues: 1-792 <CHU>  
 A/Cross-references: UNIPARC:UPI000017854F; GB:D00501  
 C/Superfamily: hepatitis C virus genome polyprotein  
 C/Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; poly  
 F/1-114/Product: capsid protein C #status predicted <CPC>  
 F/46-67/Domain: transmembrane #status predicted <TM1>  
 F/102-118/Domain: transmembrane #status predicted <TM2>  
 F/115-280/Product: membrane-associated protein M precursor #status predicted <MMP>  
 F/206-280/Product: membrane-associated protein M #status predicted <MPM>  
 F/244-263/Domain: transmembrane #status predicted <TM3>  
 F/266-281/Domain: transmembrane #status predicted <TM4>  
 F/281-775/Product: envelope protein E #status predicted <EPE>  
 F/715-735/Domain: transmembrane #status predicted <TM5>  
 F/755-773/Domain: transmembrane #status predicted <TM6>  
 F/776-792/Product: nonstructural protein NS1 (fragment) #status predicted <NS1>  
 F/183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 76.8%; Score 43; DB 2; Length 792;  
 Best Local Similarity 87.5%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 ||| |||||  
 Db 238 ETWALRHP 245

RESULT 13  
 B32401  
 genome polyprotein - dengue virus type 1 (strain AHF 82-80) (fragment)  
 N/Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru  
 C/Species: dengue virus type 1

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
 C/Accession: B32401  
 R/Chu, M.C.; O'Rourke, E.J.; Trent, D.W.  
 J. Gen. Virol. 70, 1701-1712, 1989  
 A/Title: Genetic relatedness among structural protein genes of dengue 1 virus strains.  
 A/Reference number: A32401; MUID:89293078; PMID:2738579  
 A/Accession: B32401  
 A/Molecule type: genomic RNA  
 A/Residues: 1-792 <CHU>  
 A/Cross-references: UNIPROT:P27912; UNIPARC:UPI000017854E; GB:D00501  
 C/Superfamily: hepatitis C virus genome polyprotein  
 C/Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; poly  
 F/1-114/Product: capsid protein C #status predicted <CPC>  
 F/46-67/Domain: transmembrane #status predicted <TM1>  
 F/102-118/Domain: transmembrane #status predicted <TM2>  
 F/115-280/Product: membrane-associated protein M precursor #status predicted <MMP>  
 F/206-280/Product: membrane-associated protein M #status predicted <MPM>  
 F/244-263/Domain: transmembrane #status predicted <TM3>  
 F/266-281/Domain: transmembrane #status predicted <TM4>  
 F/281-775/Product: envelope protein E #status predicted <EPE>  
 F/715-735/Domain: transmembrane #status predicted <TM5>  
 F/755-773/Domain: transmembrane #status predicted <TM6>  
 F/776-792/Product: nonstructural protein NS1 (fragment) #status predicted <NS1>  
 F/183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 76.8%; Score 43; DB 2; Length 792;  
 Best Local Similarity 87.5%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 ||| |||||  
 Db 238 ETWALRHP 245

RESULT 14  
 A32401  
 genome polyprotein - dengue virus type 1 (strain CV1636/77) (fragment)  
 N/Contains: capsid protein C; envelope protein E; membrane-associated protein M; nonstru  
 C/Species: dengue virus type 1  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
 C/Accession: A32401  
 R/Chu, M.C.; O'Rourke, E.J.; Trent, D.W.  
 J. Gen. Virol. 70, 1701-1712, 1989  
 A/Title: Genetic relatedness among structural protein genes of dengue 1 virus strains.  
 A/Reference number: A32401; MUID:89293078; PMID:2738579  
 A/Accession: A32401  
 A/Molecule type: genomic RNA  
 A/Residues: 1-792 <CHU>  
 A/Cross-references: UNIPROT:P27913; UNIPARC:UPI000017854D; GB:D00501  
 C/Superfamily: hepatitis C virus genome polyprotein  
 C/Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; poly  
 F/1-114/Product: capsid protein C #status predicted <CPC>  
 F/46-67/Domain: transmembrane #status predicted <TM1>  
 F/102-118/Domain: transmembrane #status predicted <TM2>  
 F/115-280/Product: membrane-associated protein M precursor #status predicted <MMP>  
 F/206-280/Product: membrane-associated protein M #status predicted <MPM>  
 F/244-263/Domain: transmembrane #status predicted <TM3>  
 F/266-281/Domain: transmembrane #status predicted <TM4>  
 F/281-775/Product: envelope protein E #status predicted <EPE>  
 F/715-735/Domain: transmembrane #status predicted <TM5>  
 F/755-773/Domain: transmembrane #status predicted <TM6>  
 F/776-792/Product: nonstructural protein NS1 (fragment) #status predicted <NS1>  
 F/183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 76.8%; Score 43; DB 2; Length 792;  
 Best Local Similarity 87.5%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 ||| |||||  
 Db 238 ETWALRHP 245

```

RESULT 15
GNMWVP
N;Contains: dengue virus type 1 (strain Western Pacific) (fragment)
C;Species: dengue virus type 1
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 31-Dec-2004
C;Accession: A27032
R;Mason, P.W.; McAAda, P.C.; Mason, T.L.; Fournier, M.J.
Virology 161, 262-267, 1987
A;Title: Sequence of the dengue-1 virus genome in the region encoding the three structural proteins
A;Reference number: A27032; MUID:88044504; PMID:3672932
A;Accession: A27032
A;Molecule type: Genomic RNA
A;Residues: 1-1226 <MAS>
A;Cross-references: UNIPROT:P17763; UNIPARC:UPI0000131DF1; GB:M23027; NID:g511850; PIDN:
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; nucle
F;2-114/Product: capsid protein C #status predicted <CPC>
F;43-59/Domain: transmembrane #status predicted <TM1>
F;101-117/Domain: transmembrane #status predicted <TM2>
F;115-280/Product: membrane-associated protein M precursor #status predicted <MPP>
F;115-205/Domain: nonterminal signal sequence #status predicted <SIG>
F;206-280/Product: membrane-associated protein M #status predicted <MPM>
F;468-284/Domain: transmembrane #status predicted <TM3>
F;281-775/Product: envelope protein E #status predicted <EPB>
F;384-391/Region: nucleotide-binding motif A (P-loop)
F;727-773/Domain: transmembrane #status predicted <TM4>
F;757-773/Domain: transmembrane #status predicted <TM5>
F;776-1127/Product: nonstructural protein NS1 #status predicted <NS1>
F;1128-1226/Product: nonstructural protein NS2a (fragment) #status predicted <N2A>
F;183,347,433,908,982,1190/Binding site: carbohydrate (Asn) #status predicted

Query Match 76.8%; Score 43; DB 1; Length 1226;
Best Local Similarity 87.5%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 238 ETWALRHP 245

RESULT 16
GNMWDP3
N;Contains: dengue virus type 3
a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: dengue virus type 3
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004
C;Accession: A34774
R;Osatomi, K.; Sumiyoshi, H.
Virology 176, 643-647, 1990
A;Title: Complete nucleotide sequence of dengue type 3 virus genome RNA.
A;Reference number: A34774; MUID:90266483; PMID:2345967
A;Accession: A34774
A;Molecule type: Genomic RNA
A;Residues: 1-3390 <OSA>
A;Cross-references: UNIPROT:P27915; UNIPARC:UPI0000131DFE; GB:M93130; NID:g323468; PIDN:
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
F;1-114/Product: capsid protein #status predicted <CAP>
F;46-67/Domain: transmembrane #status predicted <TM1>
F;115-280/Product: membrane protein precursor #status predicted <MBP>
F;115-205/Domain: nonterminal signal sequence #status predicted <SIG>
F;206-280/Product: membrane protein #status predicted <MEM>
F;266-280/Domain: transmembrane #status predicted <TM3>
F;281-773/Product: envelope protein #status predicted <ENV>
F;724-746/Domain: transmembrane #status predicted <TM4>
F;753-771/Domain: transmembrane #status predicted <TM5>
F;774-1184/Product: nonstructural protein NS1 #status predicted <NS1>
F;1156-1175/Domain: transmembrane #status predicted <TM6>
F;1185-1343/Product: nonstructural protein NS2a #status predicted <N2A>

Query Match 76.8%; Score 43; DB 1; Length 1226;
Best Local Similarity 87.5%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 238 ETWALRHP 245

```

```

F;1344-1473/Product: nonstructural protein NS2b #status predicted <N2B>
F;1474-2092/Product: nonstructural protein NS3 #status predicted <NS3>
F;1667-1674/Region: nucleotide-binding motif A (P-loop)
F;1754-1759/Region: nucleotide-binding motif B
F;1758-1761/Region: DEAH motif
F;2093-2378/Product: nonstructural protein NS4a #status predicted <N4A>
F;2379-2490/Product: nonstructural protein NS4b #status predicted <N4B>
F;2491-3390/Product: nonstructural protein NS5 #status predicted <NS5>
F;183,347,433,750,903,980,1132,1188,1661,2300,2304,2386,2456,2702,2712/Binding site: car
Query Match 76.8%; Score 43; DB 1; Length 3390;
Best Local Similarity 87.5%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 238 ETWALRHP 245

RESULT 17
A42551
N;Contains: dengue virus type 1 (strain Singapore S275/90)
a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: dengue virus type 1
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004
C;Accession: A42551
R;Fu, J.; Tan, B.H.; Yap, E.H.; Chan, Y.C.; Tan, Y.H.
Virology 188, 953-958, 1992
A;Title: Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90).
A;Reference number: A42551; MUID:92263809; PMID:1585663
A;Accession: A42551
A;Molecule type: Genomic RNA
A;Residues: 1-3396 <FUU>
A;Cross-references: UNIPROT:P33478; UNIPARC:UPI00002F845; GB:M87512
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
F;1-114/Product: capsid protein #status predicted <CAP>
F;115-281/Product: membrane protein precursor #status predicted <MBP>
F;115-204/Domain: nonterminal signal sequence #status predicted <SIG>
F;205-281/Product: membrane protein #status predicted <MEM>
F;267-279/Domain: transmembrane #status predicted <TM1>
F;282-774/Product: envelope protein #status predicted <ENV>
F;753-769/Domain: transmembrane #status predicted <TM2>
F;775-1127/Product: nonstructural protein NS1 #status predicted <NS1>
F;1128-1344/Product: nonstructural protein NS2a #status predicted <N2A>
F;1345-1474/Product: nonstructural protein NS2b #status predicted <N2B>
F;1475-2093/Product: nonstructural protein NS3 #status predicted <NS3>
F;1668-1675/Region: nucleotide-binding motif A (P-loop)
F;1755-1760/Region: nucleotide-binding motif B
F;1759-1762/Region: DEAH motif
F;2094-2243/Product: nonstructural protein NS4a #status predicted <N4A>
F;2244-2492/Product: nonstructural protein NS4b #status predicted <N4B>
F;2493-3396/Product: nonstructural protein NS5 #status predicted <NS5>
F;183,347,433/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 76.8%; Score 43; DB 1; Length 3396;
Best Local Similarity 87.5%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 238 ETWALRHP 245

RESULT 18
S09223
N;Contains: dengue virus type 2 (strain M1) (fragment)
C;Species: dengue virus type 2
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 31-Dec-2004
C;Accession: S09223
R;Samuel, S.; Koh, C.L.; Pang, T.; Lam, S.K.
Nucleic Acids Res. 18, 1905, 1990

```



A;Title: Nucleotide and encoded amino acid sequences of the membrane protein precursor a  
agic fever, dengue shock syndrome or dengue fever.  
A;Reference number: S09223; MUID:90245599; PMID:2336374  
A;Accession: S09223  
A;Molecule type: genomic RNA  
A;Residues: 1-166 <SAM>  
A;Cross-references: UNIPROT:Q67423; UNIPARC:UPI00000F3200; EMBL:X51713; NID:G59309; PIDN  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: membrane protein

Query Match 75.0%; Score 42; DB 2; Length 166;  
Best Local Similarity 75.0%; Pred. No. 4.2; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
: || || || ||  
Db 124 DTWILRHP 131

RESULT 19  
S09225  
membrane protein - dengue virus type 2 (strain M3) (fragment)  
C;Species: dengue virus type 2  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 31-Dec-2004  
C;Accession: S09225  
R;Samuel, S.; Koh, C.L.; Pang, T.; Lam, S.K.  
Nucleic Acids Res. 18, 1905, 1990  
A;Title: Nucleotide and encoded amino acid sequences of the membrane protein precursor a  
agic fever, dengue shock syndrome or dengue fever.  
A;Reference number: S09223; MUID:90245599; PMID:2336374  
A;Accession: S09225  
A;Molecule type: genomic RNA  
A;Residues: 1-166 <SAM>  
A;Cross-references: UNIPROT:Q67421; UNIPARC:UPI00000EFA4; EMBL:X51711; NID:G59305; PIDN  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: membrane protein

Query Match 75.0%; Score 42; DB 2; Length 166;  
Best Local Similarity 75.0%; Pred. No. 4.2; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
: || || || || ||  
Db 124 QTWILRHP 131

RESULT 20  
A83184  
probable protein methyltransferase PA3706 [imported] - Pseudomonas aeruginosa (strain PA  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: A83184  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: A83184  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-422 <STO>  
A;Cross-references: UNIPROT:Q9HXT5; UNIPARC:UPI00000CSACD; GB:AE004789; GB:AE004091; NID  
C;Genetics:  
C;Experimental source: strain PA01  
A;Gene: PA3706

Query Match 73.2%; Score 41; DB 2; Length 422;  
Best Local Similarity 75.0%; Pred. No. 16; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
: || || || || ||  
Db 124 QTWILRHP 131

RESULT 21  
S09224  
membrane protein - dengue virus type 2 (strain M2) (fragment)  
C;Species: dengue virus type 2  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 31-Dec-2004  
C;Accession: S09224  
R;Samuel, S.; Koh, C.L.; Pang, T.; Lam, S.K.  
Nucleic Acids Res. 18, 1905, 1990  
A;Title: Nucleotide and encoded amino acid sequences of the membrane protein precursor a  
agic fever, dengue shock syndrome or dengue fever.  
A;Reference number: S09223; MUID:90245599; PMID:2336374  
A;Accession: S09224  
A;Molecule type: genomic RNA  
A;Residues: 1-166 <SAM>  
A;Cross-references: UNIPROT:Q67422; UNIPARC:UPI00000F4214; EMBL:X51712; NID:G59307; PIDN  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: membrane protein

Query Match 71.4%; Score 40; DB 2; Length 166;  
Best Local Similarity 85.7%; Pred. No. 9.3; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TWFLRHP 9  
: || || || || ||  
Db 125 TWILRHP 131

RESULT 22  
E86085  
hypothetical protein Yijf [imported] - Escherichia coli (strain O157:H7, substrain EDL93  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: E86085  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: E86085  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-205 <STO>  
A;Cross-references: UNIPROT:Q8X763; UNIPARC:UPI00001659BC; GB:AE005174; NID:G12518859; P  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
C;Superfamily: Escherichia coli hypothetical 23.0K protein b3944

Query Match 69.6%; Score 39; DB 2; Length 205;  
Best Local Similarity 85.7%; Pred. No. 17; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRH 8  
: || || || || ||  
Db 126 ETWFLRH 132

RESULT 23  
A98238  
hypothetical protein ECE4873 [imported] - Escherichia coli (strain O157:H7, substrain RIN  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: A98238  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: A98238  
A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-205 <HAY>  
A;Cross-references: UNIPROT:Q8X763; UNIPARC:UPI000000A0AB9; GB:BA0000007; PIDN:BAB38296.1;  
A;Experimental source: strain O157:H7, substrain RMD 0509952  
C;Genetics:  
A;Gene: EC64873  
C;Superfamily: Escherichia coli hypothetical 23.0K protein b3944

Query Match 69.6%; Score 39; DB 2; Length 205;  
Best Local Similarity 85.7%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRH 8  
| | | | |  
Db 126 ETWFLRH 132

RESULT 24  
H95879  
probable sugar ABC transporter permease protein SMB20318 [imported] - Sinorhizobium meliloti  
C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: H95879  
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: H95879  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-343 <KUR>  
A;Cross-references: UNIPROT:Q92WM8; UNIPARC:UPI000000CB4A7; GB:AL591985; PIDN:CAC48704.1;  
A;Experimental source: strain 1021, megaplasmid pSymb  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abolia, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: SMB20318  
A;Genome: plasmid  
C;Superfamily: l-arabinose transport system permease araH

Query Match 69.6%; Score 39; DB 2; Length 343;  
Best Local Similarity 83.3%; Pred. No. 28;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 WFLRHP 9  
| | | | |  
Db 24 WFLRHP 29

RESULT 25  
PS0043  
genome polyprotein - dengue virus type 2 (strain PUO-218) (fragment)  
N;Contains: envelope protein E; membrane-associated protein M; nonstructural protein NS1  
C;Species: dengue virus type 2  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
C;Accession: PS0043  
R;Gruenberg, A.; Woo, W.S.; Biedrzycka, A.; Wright, P.J.  
J. Gen. Virol. 69, 1391-1398, 1988  
A;Title: Partial nucleotide sequence and deduced amino acid sequence of the structural p  
A;Reference number: PS0043; MUID:88258474; PMID:3385407  
A;Accession: PS0043  
A;Molecule type: mRNA  
A;Residues: 1-665 <GRU>  
A;Cross-references: UNIPROT:P18356; UNIPARC:UPI0000178550  
C;Comment: The RNA sequence was obtained from the DBJ, release 5.0.  
C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: envelope protein; glycoprotein; membrane protein; nonstructural protein; poly  
F;1-91/Domain: signal sequence #status predicted <SIG>  
F;92-166/Product: membrane-associated protein M #status predicted <MG>  
F;167-661/Product: envelope protein E #status predicted <EFE>  
F;662-665/Product: nonstructural protein NS1 (fragment) #status predicted <NS1>  
F;69,233,319/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.6%; Score 39; DB 2; Length 665;  
Best Local Similarity 75.0%; Pred. No. 55;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ETWFLRH 9  
| | | | |  
Db 124 ETWFLRH 131

RESULT 26  
I78665  
hypothetical 23.0K protein b3944 - Escherichia coli (strain K-12)  
N;Alternate names: hypothetical protein F205  
C;Species: Escherichia coli  
C;Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 09-Jul-2004  
C;Accession: I78665; C65201  
R;Blattner, F.R.; Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.  
Nucleic Acids Res. 21, 5408-5417, 1993  
A;Title: Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89  
A;Reference number: I58303; MUID:94089392; PMID:8265357  
A;Accession: I78665  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-205 <RES>  
A;Cross-references: UNIPROT:P32668; UNIPARC:UPI000013B429; EMBL:U00006; NID:q409785; PIDN:  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: C65201  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-205 <BLAT>  
A;Cross-references: UNIPARC:UPI000013B429; GB:AE000468; GB:U00096; NID:gl790374; PIDN:ACC  
A;Experimental source: strain K-12, substrain MG1655  
C;Genetics:  
A;Gene: yijF  
C;Superfamily: Escherichia coli hypothetical 23.0K protein b3944

Query Match 67.9%; Score 38; DB 2; Length 205;  
Best Local Similarity 85.7%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRH 8  
| | | | |  
Db 126 ETWFLRH 132

RESULT 27  
I41293  
EcoS type I restriction modification enzyme M subunit - Escherichia coli  
C;Species: Escherichia coli  
C;Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004  
C;Accession: I41293  
R;Murray, N.E.; Daniel, A.S.; Cowan, G.M.; Sharp, P.M.  
Mol. Microbiol. 9, 133-143, 1993  
A;Title: Conservation of motifs within the unusually variable polypeptide sequences of ty  
A;Reference number: I41293; MUID:94018600; PMID:8412658  
A;Accession: I41293  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-490 <RES>  
A;Cross-references: UNIPROT:Q47282; UNIPARC:UPI0000136789; GB:L18759; NID:g304895; PIDN:  
C;Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match 67.9%; Score 38; DB 2; Length 490;  
 Best Local Similarity 55.6%; Pred. No. 60;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RETWFLRHP 9  
 Db 374 KEVWFYHP 382

RESULT 28  
 B96761  
 Probable protein kinase T9L24.36 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: B96761  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: B96761  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1155 <STO>  
 A:Cross-references: UNIPROT:Q9FX38; UNIPARC:UPI000009F5D2; GB:AE005173; NID:g11120796; F  
 C:Genetics:  
 A:Gene: T9L24.36  
 A:Map position: 1

Query Match 67.9%; Score 38; DB 2; Length 1155;  
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 8  
 Db 433 ETWFLAH 439

RESULT 29  
 S49184  
 Phosphinothricin N-acetyltransferase (EC 2.3.1.1) - Streptomyces griseus  
 C:Species: Streptomyces griseus  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: S49184  
 R:Marcos, A.T.; Diez, B.; Gutierrez, S.; Fernandez, F.J.; Oguiza, J.A.; Martin, J.F.  
 submitted to the EMBL Data Library, June 1994  
 A:Description: Three genes hrdB, hrdD and hrdT of Streptomyces griseus IMRU 3570, encodi  
 A:Reference number: S49183  
 A:Accession: S49184  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-194 <MAR>  
 A:Cross-references: UNIPROT:Q54225; UNIPARC:UPI000012FDA5; EMBL:X79980; NID:g510451; PID  
 C:Superfamily: phosphinothricin N-acetyltransferase  
 C:Keywords: acyltransferase

Query Match 66.1%; Score 37; DB 1; Length 194;  
 Best Local Similarity 66.7%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RETWFLRHP 9  
 Db 47 RMQWFLSHP 55

RESULT 30  
 C95872

hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymB  
 C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
 C:Accession: C95872  
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo  
 A:Reference number: A95842; MUID:21396508; PMID:11481431  
 A:Accession: C95872  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-301 <KUR>  
 A:Cross-references: UNIPROT:Q92WT6; UNIPARC:UPI00000CB46E; GB:AL591985; PIDN:CAC48643.1;  
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler  
 dela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: SMB20253  
 A:Genome: plasmid

Query Match 66.1%; Score 37; DB 2; Length 301;  
 Best Local Similarity 75.0%; Pred. No. 55;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 Db 57 EVMNLRHP 64

Search completed: August 31, 2006, 11:51:50  
 Job time : 19.25 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein.search, using sw model  
Run on: August 31, 2006, 11:33:43 ; Search time 139 Seconds  
(without alignments)  
59.893 Million cell updates/sec

Title: DENGUE\_SEROTYPE3  
Perfect score: 56  
Sequence: 1 retwflrhp 9  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2849598 seqs, 925015592 residues  
Total number of hits satisfying chosen parameters: 2849598  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_7.2.1\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	47	83.9	4112 2	Q57V78	9TRYTP
2	45	80.4	120 2	Q67424	dengue viru
3	45	80.4	166 2	Q66346	dengue viru
4	45	80.4	166 2	Q66357	dengue viru
5	45	80.4	280 2	Q63264	dengue viru
6	45	80.4	280 2	Q63265	dengue viru
7	45	80.4	280 2	Q63266	dengue viru
8	45	80.4	280 2	Q63267	dengue viru
9	45	80.4	555 1	POLG_DEN2H	
10	45	80.4	565 2	Q32P70	9FLAV
11	45	80.4	578 2	Q12290	9FLAV
12	45	80.4	661 2	Q3BCV3	9FLAV
13	45	80.4	661 2	Q3BCV4	9FLAV
14	45	80.4	661 2	Q3BCV5	9FLAV
15	45	80.4	661 2	Q3BCX6	dengue viru
16	45	80.4	661 2	Q3BCX7	9FLAV
17	45	80.4	661 2	Q3BCX8	9FLAV
18	45	80.4	661 2	Q3BCX9	9FLAV
19	45	80.4	661 2	Q3BCY0	dengue viru
20	45	80.4	661 2	Q3BCY1	dengue viru
21	45	80.4	661 2	Q3BCY2	dengue viru
22	45	80.4	661 2	Q3BCY3	dengue viru
23	45	80.4	661 2	Q3BCY4	dengue viru
24	45	80.4	661 2	Q3BCY5	dengue viru
25	45	80.4	661 2	Q5QIB6	9FLAV
26	45	80.4	661 2	Q5QIB7	dengue viru
27	45	80.4	661 2	Q5QIB8	dengue viru
28	45	80.4	661 2	Q5QIB9	9FLAV
29	45	80.4	661 2	Q5QIB0	9FLAV
30	45	80.4	661 2	Q5QIB1	9FLAV
31	45	80.4	661 2	Q5QIB2	dengue viru

32	45	80.4	661	2	Q5VI93	9FLAV	Q5VI93	dengue viru
33	45	80.4	661	2	Q5VI94	9FLAV	Q5VI94	dengue viru
34	45	80.4	661	2	Q5VI95	9FLAV	Q5VI95	dengue viru
35	45	80.4	661	2	Q5VI96	9FLAV	Q5VI96	dengue viru
36	45	80.4	716	2	Q6DUV2	9FLAV	Q6DUV2	dengue viru
37	45	80.4	724	2	Q5ICU9	9FLAV	Q5ICU9	dengue viru
38	45	80.4	745	2	Q6KEK9	9FLAV	Q6KEK9	dengue viru
39	45	80.4	757	2	Q5S8P1	9FLAV	Q5S8P1	dengue viru
40	45	80.4	757	2	Q5S8P2	9FLAV	Q5S8P2	dengue viru
41	45	80.4	757	2	Q6DUD9	9FLAV	Q6DUD9	dengue viru
42	45	80.4	763	2	Q5ICU8	9FLAV	Q5ICU8	dengue viru
43	45	80.4	775	2	Q66398	9FLAV	Q66398	dengue viru
44	45	80.4	775	2	Q8QY07	9FLAV	Q8QY07	dengue viru
45	45	80.4	775	2	Q8QY62	9FLAV	Q8QY62	dengue viru
46	45	80.4	775	2	Q8QY63	9FLAV	Q8QY63	dengue viru
47	45	80.4	779	2	Q88636	9FLAV	Q88636	dengue viru
48	45	80.4	1127	1	POLG_DEN2D		P30026	dengue viru
49	45	80.4	1127	2	P87638	9FLAV	P87638	dengue viru
50	45	80.4	1127	2	P89531	9FLAV	P89531	dengue viru
51	45	80.4	1127	2	P89532	9FLAV	P89532	dengue viru
52	45	80.4	1127	2	Q66454	9FLAV	Q66454	dengue viru
53	45	80.4	1127	2	Q66455	9FLAV	Q66455	dengue viru
54	45	80.4	1127	2	Q66456	9FLAV	Q66456	dengue viru
55	45	80.4	1127	2	Q66457	9FLAV	Q66457	dengue viru
56	45	80.4	3388	1	POLG_DEN2P		P12823	d genome po
57	45	80.4	3391	1	POLG_DEN26		P29990	d genome po
58	45	80.4	3391	1	POLG_DEN27		P29991	d genome po
59	45	80.4	3391	1	POLG_DEN2J		P07564	d genome po
60	45	80.4	3391	1	POLG_DEN2N		P14340	d genome po
61	45	80.4	3391	1	Q09234	DEN26	Q09234	dengue viru
62	45	80.4	3391	2	Q11875	9FLAV	Q11875	dengue viru
63	45	80.4	3391	2	Q92752	9FLAV	Q92752	dengue viru
64	45	80.4	3391	2	Q92753	9FLAV	Q92753	dengue viru
65	45	80.4	3391	2	Q92754	9FLAV	Q92754	dengue viru
66	45	80.4	3391	2	Q92835	9FLAV	Q92835	dengue viru
67	45	80.4	3391	2	Q58Y66	9FLAV	Q58Y66	dengue viru
68	45	80.4	3391	2	Q58Y67	9FLAV	Q58Y67	dengue viru
69	45	80.4	3391	2	Q58Y69	9FLAV	Q58Y69	dengue viru
70	45	80.4	3391	2	Q58Y71	9FLAV	Q58Y71	dengue viru
71	45	80.4	3391	2	Q58Y71	9FLAV	Q58Y71	dengue viru
72	45	80.4	3391	2	Q58Y71	9FLAV	Q58Y71	dengue viru
73	45	80.4	3391	2	Q68Y26	9FLAV	Q68Y26	dengue viru
74	45	80.4	3391	2	Q70YQ7	9FLAV	Q70YQ7	dengue viru
75	45	80.4	3391	2	Q8QR27	9FLAV	Q8QR27	dengue viru
76	45	80.4	3391	2	Q91SD1	9FLAV	Q91SD1	dengue viru
77	45	80.4	3391	2	Q91U94	9FLAV	Q91U94	dengue viru
78	45	80.4	3391	2	Q9E7P0	9FLAV	Q9E7P0	dengue viru
79	45	80.4	3391	2	Q91F59	9FLAV	Q91F59	dengue viru
80	45	80.4	3391	2	Q9J8D1	9FLAV	Q9J8D1	dengue viru
81	45	80.4	3391	2	Q9J8D2	9FLAV	Q9J8D2	dengue viru
82	45	80.4	3391	2	Q9J8D3	9FLAV	Q9J8D3	dengue viru
83	45	80.4	3391	2	Q9J8D4	9FLAV	Q9J8D4	dengue viru
84	45	80.4	3391	2	Q9J8D5	9FLAV	Q9J8D5	dengue viru
85	45	80.4	3391	2	Q9J8D6	9FLAV	Q9J8D6	dengue viru
86	45	80.4	3391	2	Q9J8D7	9FLAV	Q9J8D7	dengue viru
87	45	80.4	3391	2	Q9J8D8	9FLAV	Q9J8D8	dengue viru
88	45	80.4	3391	2	Q9J8D9	9FLAV	Q9J8D9	dengue viru
89	45	80.4	3391	2	Q9J8E0	9FLAV	Q9J8E0	dengue viru
90	45	80.4	3391	2	Q9J8E1	9FLAV	Q9J8E1	dengue viru
91	45	80.4	3391	2	Q9Q4T1	9FLAV	Q9Q4T1	dengue viru
92	45	80.4	3391	2	Q9Q4T2	9FLAV	Q9Q4T2	dengue viru
93	45	80.4	3391	2	Q9W8I3	9FLAV	Q9W8I3	dengue viru
94	45	80.4	3391	2	Q9WD99	9FLAV	Q9WD99	dengue viru
95	45	80.4	3391	2	Q9WDA0	9FLAV	Q9WDA0	dengue viru
96	45	80.4	3391	2	Q9WDA1	9FLAV	Q9WDA1	dengue viru
97	45	80.4	3391	2	Q9WDA2	9FLAV	Q9WDA2	dengue viru
98	45	80.4	3391	2	Q9WDA3	9FLAV	Q9WDA3	dengue viru
99	45	80.4	3391	2	Q9WDA4	9FLAV	Q9WDA4	dengue viru
100	45	80.4	3391	2	Q9WDA5	9FLAV	Q9WDA5	dengue viru
					Q9WDA6	9FLAV	Q9WDA6	dengue viru

ALIGNMENTS

```

RESULT 1
Q57V78_9TRYR PRELIMINARY; PRT; 4112 AA.
AC Q57V78;
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Dynein heavy chain, putative.
GN ORFNames=TB927.7.920;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUT10.1;
RA Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Shalton J., Hou L., Dijkeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AC159437; AAX70281.1; -; Genomic_DNA.
DR GO; GO:0030286; C:dynactin complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0003777; F:microtubule motor activity; IEA.
DR GO; GO:0007018; P:microtubule-based movement; IEA.
DR InterPro; IPR011704; AAA_5.
DR InterPro; IPR004273; Dynein_heavy.
DR Pfam; PF07728; AAA_5; 1.
DR Pfam; PF03028; Dynein_heavy; 1.
SQ SEQUENCE 4112 AA; 468087 MW; EC33607D9855EDA0 CRC64;

Query Match 83.9%; Score 47; DB 2; Length 4112;
Best Local Similarity 77.8%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RETWFLRHP 9
Db 1240 RETWFFNHP 1248

RESULT 2
Q67424_9FLAV PRELIMINARY; PRT; 120 AA.
AC Q67424;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Genomic RNA for envelope protein E N-term. (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=New Guinea C;
RA MEDLINE=87197230; PubMed=2952760;
RA Biedrzycka A., Cauchi M.R., Bartholomeusz A., Gorman J.J.,
RA Wright P.J.;
RT "Characterization of protease cleavage sites involved in the formation
RT of the envelope glycoprotein and three non-structural proteins of
RT dengue virus type 2, New Guinea C strain.";
RL J. Gen. Virol. 68:1317-1326(1987).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

DR EMBL; X05375; CAA28966.1; -; Genomic_RNA.
DR HSP; O88653; 10KE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Envelope protein.
FT CHAIN 18 92 protein M.
FT CHAIN 93 >120 protein E.
FT NON_TER 1
FT NON_TER 120
SQ SEQUENCE 120 AA; 13329 MW; FF86913787CA5C27 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 120;
Best Local Similarity 87.5%; Pred. No. 6.2;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
Db 50 ETWILRHP 57

RESULT 3
Q66346_9FLAV PRELIMINARY; PRT; 166 AA.
AC Q66346;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Premembrane polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TH-36;
RA Shiu S.Y.W.;
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; X72849; CAA51363.1; -; mRNA.
DR PIR; S40144; S40144.
DR GO; GO:0019028; P:viral capsid; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT CHAIN 92 >166 membrane protein.
FT NON_TER 1
FT NON_TER 166
SQ SEQUENCE 166 AA; 18751 MW; F498748D35909639 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
Db 124 ETWILRHP 131

RESULT 4
Q66357_9FLAV
```

```

ID Q6357_9FLAV PRELIMINARY; PRT; 166 AA.
AC Q6357;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Membrane protein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D2-04;
RX MEDLINE=94269837; PubMed=7911607;
RA Yang P.Y., Kautner I., Koh C.L., Lam S.K.;
RT "Nucleotide and deduced amino acid sequences of genes encoding the
RT structural and nonstructural NS1 proteins of a dengue-2 virus isolated
RT in China.";
RL Virus Genes 8:71-74 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D2-04;
RA Koh C.;
RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; X65242; CAA46343.1; -; Genomic_RNA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
FT CHAIN 92 >166 membrane protein precursor.
FT NON_TER 1
FT NON_TER 166
SQ SEQUENCE 166 AA; 18790 MW; 07E1FC1EBB7D4521 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 124 ETWILRHP 131

RESULT 5
Q8QZ64_9FLAV PRELIMINARY; PRT; 280 AA.
AC Q8QZ64;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21571640; PubMed=11714970;
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,
RA Holmes E.C., Gould E.A.;
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT for in situ virus evolution and recombination.";
RL J. Gen. Virol. 82:2945-2953(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

```

```

DR EMBL; AF360863; AAL76291.1; -; Genomic_RNA.
DR SMR; Q8QZ64; 21-100.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR001122; Flavi_capsidC.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 280
SQ SEQUENCE 280 AA; 31847 MW; E889FDD11929CBA7 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 280;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 238 ETWILRHP 245

RESULT 6
Q8QZ65_9FLAV PRELIMINARY; PRT; 280 AA.
AC Q8QZ65;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21571640; PubMed=11714970;
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,
RA Holmes E.C., Gould E.A.;
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT for in situ virus evolution and recombination.";
RL J. Gen. Virol. 82:2945-2953(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF360862; AAL76290.1; -; Genomic_RNA.
DR SMR; Q8QZ65; 21-100.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR001122; Flavi_capsidC.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 280
SQ SEQUENCE 280 AA; 31849 MW; BEB9F24A8D29CFBD CRC64;

Query Match 80.4%; Score 45; DB 2; Length 280;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 238 ETWILRHP 245

RESULT 7

```

```
Q8QZ66_9FLAV
ID Q8QZ66_9FLAV PRELIMINARY; PRT; 280 AA.
AC Q8QZ66;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21571640; PubMed=11714970;
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,
RA Holmes E.C., Gould E.A.;
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT for in situ virus evolution and recombination.";
RL J. Gen. Virol. 82:2945-2953(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF360861; AAL76289.1; -; Genomic_RNA.
DR SMR; Q8QZ66; 21-100.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR001122; Flavi_capsidC.
DR InterPro; IPR000069; Flavi_M.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
DR Polyprotein.
DR NON_TER 280 280
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21571640; PubMed=11714970;
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,
RA Holmes E.C., Gould E.A.;
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT for in situ virus evolution and recombination.";
RL J. Gen. Virol. 82:2945-2953(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF360861; AAL76289.1; -; Genomic_RNA.
DR SMR; Q8QZ66; 21-100.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR001122; Flavi_capsidC.
DR InterPro; IPR000069; Flavi_M.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
DR Polyprotein.
DR NON_TER 280 280
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=21571640; PubMed=11714970;
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,
RA Holmes E.C., Gould E.A.;
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT for in situ virus evolution and recombination.";
RL J. Gen. Virol. 82:2945-2953(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF360860; AAL76288.1; -; Genomic_RNA.
DR SMR; Q8QZ67; 21-100.
DR
```

```
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR001122; Flavi_capsidC.
DR InterPro; IPR000069; Flavi_M.
DR Pfam; PF01003; Flavi_capsid; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
DR Polyprotein.
DR NON_TER 280 280
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=21571640; PubMed=11714970;
RA Uzcategui N.Y., Camacho D., Comach G., Cuello de Uzcategui R.,
RA Holmes E.C., Gould E.A.;
RT "Molecular epidemiology of dengue type 2 virus in Venezuela: evidence
RT for in situ virus evolution and recombination.";
RL J. Gen. Virol. 82:2945-2953(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF360860; AAL76288.1; -; Genomic_RNA.
DR SMR; Q8QZ67; 21-100.
DR
```

```
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0005198; P:structural molecule activity; IEA.
GO; GO:0019058; P:viral infectious cycle; IEA.
InterPro; IPR001122; Flavi_capsidC.
InterPro; IPR000069; Flavi_M.
Pfam; PF01003; Flavi_capsid; 1.
Pfam; PF01004; Flavi_M; 1.
Pfam; PF01570; Flavi_propep; 1.
Polyprotein.
NON_TER 280 280
NCBI_TaxID=11060;
SEQUENCE 280 AA; 31893 MW; 814A8B9B4A22AC20 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 280;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
   ||| ||||
Db 238 ETWILRHP 245

RESULT 9
POLG DEN2H
ID POLG DEN2H STANDARD; PRT; 555 AA.
AC P29984;
DT 01-APR-1993, integrated into UniProtKB/Swiss-Prot.
DT 01-APR-1993, sequence version 1.
DT 01-APR-2006, entry version 43.
DE Genome polyprotein [Contains: Envelope protein M (Matrix protein);
DE Major envelope protein E; Nonstructural protein 1 (NS1)] (Fragment).
OS Dengue virus type 2 (strain TH-36).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=31637;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=92113574; PubMed=1339466;
RA Shiu S.Y.W., Jiang W.R., Porterfield J.S., Gould E.A.;
RT "Envelope protein sequences of dengue virus isolates TH-36 and TH-
RT Swan, and identification of a type-specific genetic marker for dengue
RT and tick-borne flaviviruses.";
RL J. Gen. Virol. 73:2027-212(1992).
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins
CC (By similarity).
CC -1- MISCELLANEOUS: The virion of this virus is a nucleocapsid covered
CC by a lipoprotein envelope. The envelope contains two proteins: the
CC protein M and glycoprotein E. The nucleocapsid is a complex of
CC protein C and mRNA. In immature particles, there are 60
CC icosahedrally organized trimeric spikes on the surface. Each spike
CC consists of three heterodimers of envelope protein M precursor
CC (prM) and envelope protein E (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; D10514; BAA01389.1; -; Genomic_RNA.
DR FRR; JQ1404; JQ1404.
DR HSP; Q88653; IOKS.
DR SMR; P29984; 50-443.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycop_C; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Capsid protein; Core protein; Envelope protein; Glycoprotein;
DR Membrane; Polyprotein; Structural protein; Transmembrane.
FT CHAIN <1 49 Envelope protein M.
FT CHAIN 50 544 /FTID=PRO_0000037922
FT CHAIN 50 544 Major envelope protein E.
FT /FTID=PRO_0000037923.
```



FT CHAIN 545 >555 Nonstructural protein 1.  
 /FTID=PRO\_0000037924.  
 FT TRANSMEM 37 53 Potential.  
 FT TRANSMEM 496 512 Potential.  
 FT TRANSMEM 526 542 Potential.  
 FT CARBOHYD 116 116 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 202 202 N-linked (GlcNAc...) (Potential).  
 FT DISULFID 52 79 By similarity.  
 FT DISULFID 109 170 By similarity.  
 FT DISULFID 123 154 By similarity.  
 FT DISULFID 141 165 By similarity.  
 FT DISULFID 234 334 By similarity.  
 FT DISULFID 351 382 By similarity.  
 FT NON\_TER 1 1  
 FT NON\_TER 555  
 SQ SEQUENCE 555 AA; 61243 MW; F8DEA740BB4DBD8DF CRC64;

Query Match 80.4%; Score 45; DB 1; Length 555;  
 Best Local Similarity 87.5%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 ||| ||||  
 Db 7 ETWILRHP 14

RESULT 10  
 ID Q3ZPJ0\_9FLAV PRELIMINARY; PRT; 565 AA.  
 AC Q3ZPJ07.  
 DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.  
 DT 27-SEP-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 3.  
 DE Polypeptide (Fragment).  
 DE Dengue virus type 2.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus; Dengue virus group.  
 OX NCBI\_TaxID=11060;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=3T90.  
 RA Pyke A.T., Hanna J., Richards A., Taylor C.T., Morgan A.,  
 RA Humphreys J., Brookes D., Smith G.A.;  
 RT "Defining Dengue in the New Millennium."  
 RL Arbovirus Res. Aust. 0:0-0(2005).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AY706011; AAWG2469.1; -; Genomic RNA.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0019058; P:viral infectious cycle; IEA.  
 DR InterPro; IPR000069; Flavi\_M.  
 DR InterPro; IPR001157; Flavi\_NSI.  
 DR InterPro; IPR00336; Flv\_glyc Ig-like.  
 DR Pfam; PF02832; Flavi\_glycop\_C; 1.  
 DR Pfam; PF00869; Flavi\_glycop\_C; 1.  
 DR Pfam; PF01004; Flavi\_M; 1.  
 DR Pfam; PF00948; Flavi\_NSI; 1.  
 DR ProDom; PD001496; Flavi\_NSI; 1.  
 DR Polyprotein.  
 KW POLYPEPTIDE.

FT CHAIN <1 47 membrane protein.  
 FT CHAIN 48 542 envelope protein.  
 FT CHAIN 543 >578 nonstructural protein 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 578 578  
 SQ SEQUENCE 565 AA; 61930 MW; 17DC94EEC53B3EF6 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 565;  
 Best Local Similarity 87.5%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

FT NON\_TER 1 1  
 FT NON\_TER 565 565  
 SQ SEQUENCE 565 AA; 61930 MW; 17DC94EEC53B3EF6 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 565;  
 Best Local Similarity 87.5%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 ||| ||||  
 Db 3 ETWILRHP 10

RESULT 11  
 ID Q12290\_9FLAV PRELIMINARY; PRT; 578 AA.  
 AC Q122907.  
 DT 01-JUL-1997, integrated into UniProtKB/TrEMBL.  
 DT 01-JUL-1997, sequence version 1.  
 DT 07-FEB-2006, entry version 26.  
 DE Polypeptide (Fragment).  
 DE Dengue virus type 2.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus; Dengue virus group.  
 OX NCBI\_TaxID=11060;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Torres Strait 1;  
 RA Serafin I.L., Phillips D.A.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AF004019; AAB61366.1; -; mRNA.  
 DR HSP; Q88653; IOKE.  
 DR SMR; Q12290; 48-441.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0019028; C:viral capsid; IEA.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0019058; P:viral infectious cycle; IEA.  
 DR InterPro; IPR001157; Flavi\_NSI.  
 DR InterPro; IPR00336; Flv\_glyc Ig-like.  
 DR InterPro; IPR011998; Vrl\_glyc\_cen\_dim.  
 DR Pfam; PF02832; Flavi\_glycop\_C; 1.  
 DR Pfam; PF00869; Flavi\_glycop\_C; 1.  
 DR Pfam; PF01004; Flavi\_M; 1.  
 DR Pfam; PF00948; Flavi\_NSI; 1.  
 DR ProDom; PD001496; Flavi\_NSI; 1.  
 DR Polyprotein.  
 KW POLYPEPTIDE.

FT CHAIN <1 47 membrane protein.  
 FT CHAIN 48 542 envelope protein.  
 FT CHAIN 543 >578 nonstructural protein 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 578 578  
 SQ SEQUENCE 578 AA; 63606 MW; 1C03A7CFD72C567D CRC64;

Query Match 80.4%; Score 45; DB 2; Length 578;  
 Best Local Similarity 87.5%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9  
 ||| ||||  
 Db 5 ETWILRHP 12

RESULT 12  
 ID Q3BCV3\_9FLAV PRELIMINARY; PRT; 661 AA.  
 AC Q3BCV37.  
 DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.  
 DT 22-NOV-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 3.  
 DE Polypeptide (Fragment).  
 DE Dengue virus type 2.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus; Dengue virus group.

Query Match 80.4%; Score 45; DB 2; Length 578;  
 Best Local Similarity 87.5%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=01st193/2001;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RL GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RN Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=01st193/2001;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786398; AAX18216.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73049 MW; 2A644DEADA728CF2 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131

RESULT 13
Q3BCV4_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCV4;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00U3/2000/human;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RL GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RN Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00U3/2000/human;
RX PubMed=16222028;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786397; AAX18215.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.

OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00U3/2000/human;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RL GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RN Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00U3/2000/human;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786399; AAX18215.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73081 MW; 6F6C51D6BEC33CA8 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131

RESULT 15
Q3BCX6_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX6;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.

DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73081 MW; 6F6C51D6BEC33CA8 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131

RESULT 15
Q3BCX6_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX6;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.

DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73072 MW; 654A28D6B96FBSA8 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131

RESULT 14
Q3BCV5_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCV5;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SLMC70/1995/human;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RL GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RN Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SLMC70/1995/human;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786396; AAX18214.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73081 MW; 6F6C51D6BEC33CA8 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131
```

```

DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St23/2000;
RX Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St23/2000;
RX Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786375; AAX18193.1; -; Genomic_RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polypeptide.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73143 MW; 5509B5931AE2BF2B CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131

RESULT 16
Q3BCX7_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX7;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St22/2000;
RX Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St22/2000;
RX Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St23/2000;
RX Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=00St23/2000;
RX Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

```

```

DR EMBL; AY786374; AAX18192.1; -; Genomic_RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polypeptide.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73083 MW; 5F56106DA1550EF6 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131

RESULT 17
Q3BCX8_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCX8;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99Sa695/1999;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=99Sa695/1999;
RX Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY786373; AAX18191.1; -; Genomic_RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polypeptide.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73086 MW; 899A28D6B96FB5B0 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131

RESULT 18
Q3BCX9_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCX9_9FLAV

```

```
AC Q3BCX9;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCMC60/1998;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCMC60/1998;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY786372; AAX18190.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73166 MW; 84C50AFD2358F08C CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
Db 124 ETWILRHP 131

RESULT 20
Q3BCYL_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCYL1;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CSMC7/1996;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CSMC7/1996;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY786370; AAX18188.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73054 MW; 751344A7E73C46F CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ETWFLRHP 9
Db 124 ETWILRHP 131

RESULT 19
Q3BCY0_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY0;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C127/1998;
RX PubMed=16222028;
RA Salda L.T., Parquet M.D., Matias R.R., Natividad F.F., Kobayashi N.,
RA Morita K.;
RT "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C127/1998;
RA Salda L.T.D.;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
```

```
RESULT 21
Q3BCY2_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCY2_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY2;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRL3/1996;
RX PubMed=16222028;
RA Salda L.T.D.;
RA Morita K.;
RA "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRL3/1996;
RA Salda L.T.D.;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY786369; AAX18187.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON TER 1 661
FT NON TER 661
SQ SEQUENCE 661 AA; 73072 MW; 654A28D6B96FB5A8 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 124 ETWILRHP 131

RESULT 22
Q3BCY3_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCY3_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY3;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH97/1995;
RX PubMed=16222028;
RA Salda L.T.;
RA Morita K.;
RA "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH90/1995;
RA Salda L.T.D.;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY786367; AAX18185.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON TER 1 661
FT NON TER 661
SQ SEQUENCE 661 AA; 73150 MW; 654F225FA969639F CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 124 ETWILRHP 131

RESULT 23
Q3BCY4_9FLAV PRELIMINARY; PRT; 661 AA.
ID Q3BCY4_9FLAV PRELIMINARY; PRT; 661 AA.
AC Q3BCY4;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH90/1995;
RX PubMed=16222028;
RA Salda L.T.;
RA Parquet M.D.;
RA Morita K.;
RA "MOLECULAR EPIDEMIOLOGY OF DENGUE 2 VIRUSES IN THE PHILIPPINES:
RT GENOTYPE SHIFT AND LOCAL EVOLUTION.";
RL Am. J. Trop. Med. Hyg. 73:796-802(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DOH90/1995;
RA Salda L.T.D.;
RA Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; AY786368; AAX18186.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
KW Polyprotein.
FT NON TER 1 661
FT NON TER 661
SQ SEQUENCE 661 AA; 73096 MW; CF865AAE54ADE0F1 CRC64;

Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 124 ETWILRHP 131
```



```

DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73119 MW; CE2051C17F40A623 CRC64;

Query Match      80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 124 ETWILRHP 131

RESULT 27
QSVI88_9FLAV
ID QSVI88_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI88;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=13381/Chochola 02;
RX PubMed=15516647;
RA Loroño-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY449683; AAS14974.1; -; Genomic_RNA.
DR SMR; QSVI88; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;

Query Match      80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ETWFLRHP 9
Db 124 ETWILRHP 131

RESULT 29
QSVI90_9FLAV
ID QSVI90_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI90;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polyprotein (Fragment).
OS Dengue virus type 2.

```

```
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=11936/St. Elena 01;
RX PubMed=15516647;
RA Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY449681; AAS14972.1; -; Genomic_RNA.
DR SMR; QSVI90; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flav_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;
Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131
RESULT 30
QSVI91_9FLAV PRELIMINARY; PRT; 661 AA.
AC QSVI91;
DT 07-DEC-2004, integrated into UniprotKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Polypeptin (Fragment).
OS Dengue virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11060;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=12914/Tekax 01;
RX PubMed=15516647;
RA Lorono-Pino M.A., Farfan-Ale J.A., Zapata-Peraza A.L.,
RA Rosado-Paredes E.P., Flores-Flores L.F., Garcia-Rejon J., Diaz F.J.,
RA Blitvich B.J., Jimenez-Rios E., Blair C.D., Olson K.E., Black W. IV,
RA Beaty B.J.;
RT "Introduction of the American/Asian genotype of dengue 2 virus into
RT the Yucatan State of Mexico.";
RL Am. J. Trop. Med. Hyg. 71:485-492(2004).
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
```

```
CC -----
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY449680; AAS14971.1; -; Genomic_RNA.
DR SMR; QSVI91; 167-560.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flav_M.
DR InterPro; IPR002535; Flavi_propep.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 73080 MW; 5216054D684173C0 CRC64;
Query Match 80.4%; Score 45; DB 2; Length 661;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ETWFLRHP 9
DB 124 ETWILRHP 131
Search completed: August 31, 2006, 11:43:08
Job time : 142 secs
```



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 11:33:43 ; Search time 110.25 Seconds  
(without alignments)  
37.324 Million cell updates/sec

Title: DENGUE\_SEROTYPE4

Perfect score: 52

Sequence: 1 veswflnp 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_8.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

9: Geneseqp2005s.\*

10: Geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	88.5	167	8	ADN37495 Dengue vi
2	46	88.5	167	8	ADN37499 Dengue vi
3	46	88.5	635	2	AAW75412 Fusion pr
4	46	88.5	675	8	ADN37625 DEN-4/Den
5	46	88.5	675	8	ADN37627 DEN-4/Den
6	46	88.5	675	8	ADN37514 Dengue vi
7	46	88.5	675	8	ADN37520 Dengue vi
8	46	88.5	675	8	ADN37516 Dengue vi
9	46	88.5	677	8	ADN37605 Dengue vi
10	46	88.5	677	8	ADN37718 Dengue vi
11	46	88.5	677	8	ADN37720 Dengue vi
12	46	88.5	678	8	ADN37623 Dengue vi
13	46	88.5	725	2	AAW06594 Amino aci
14	46	88.5	3387	4	AAE07991 Wild-type
15	46	88.5	3387	4	AAE07992 Attenuate
16	46	88.5	3387	6	AAE35313 Recombina
17	46	88.5	3387	6	AAE35312 Dengue vi
18	46	88.5	3391	4	AAE07985 Dengue vi
19	43	82.7	39	9	ADW12582 DEN-4/DEN
20	43	82.7	40	9	ADW12577 M1-40/YF.
21	43	82.7	48	9	ADW12588 p(95-114)
22	43	82.7	680	8	ADR87182 Yellow Fe
23	43	82.7	681	6	ABP57863 Plasmid p

24	43	82.7	681	6	ABP57861	Abp57861 Plasmid p
25	43	82.7	3411	8	ADJ57394	Adj57394 Hamster p
26	42	80.8	167	9	ADM00905	Adm00905 Amino aci
27	42	80.8	668	9	ADM00902	Adm00902 Amino aci
28	42	80.8	691	8	ADS76177	AdS76177 Heterodim
29	42	80.8	692	8	ABP57859	Abp57859 Plasmid p
30	42	80.8	692	6	ABP57862	Abp57862 Plasmid p
31	42	80.8	694	8	ADR87178	AdR87178 West Nile
32	42	80.8	3402	4	AAE07988	Aae07988 Dengue vi
33	42	80.8	3430	7	ADJ92005	Adj92005 West Nile
34	42	80.8	3430	8	ADK13682	Adk13682 West Nile
35	42	80.8	3433	6	ABB98821	Abb98821 West Nile
36	42	80.8	3433	6	ABP70647	Abp70647 Amino aci
37	42	80.8	3433	9	ADM00898	Adm00898 Amino aci
38	42	80.8	3433	9	AEBA4329	Aeb44329 West Nile
39	42	80.8	3433	9	AEBA4319	Aeb44319 West Nile
40	41	78.8	59	5	ADT77245	Adt77245 Human alp
41	41	78.8	60	5	ABB84465	Abb84465 alpha2 ho
42	39	75.0	27	8	ADN11217	Adn11217 Peptide m
43	39	75.0	27	8	ADN11193	Adn11193 Peptide m
44	39	75.0	619	8	ADR88954	Adr88954 A. thalia
45	39	75.0	1197	4	AAE92739	Aag92739 C. glutami
46	39	75.0	1197	9	AED71804	Aed71804 Corynebac
47	38	73.1	69	4	AAW92279	Aam92279 Human dig
48	38	73.1	278	8	ADQ25888	Adq25888 Human GPC
49	38	73.1	333	3	AAG41837	Aag41837 Arabidops
50	38	73.1	348	3	AAG41836	Aag41836 Arabidops
51	38	73.1	386	3	AAG41835	Aag41835 Arabidops
52	38	73.1	386	5	ABB93987	Abb93987 Herbicida
53	38	73.1	455	6	AAO31014	Aao31014 Human tra
54	38	73.1	463	5	ABP74118	Abp74118 Human TRI
55	38	73.1	546	9	AEBA1909	Aeb41909 L. pneumo
56	38	73.1	557	9	AEBA38677	Aeb38677 L. pneumo
57	38	73.1	669	4	ABB69903	Abb69903 Drosophil
58	38	73.1	826	5	ABB07253	Abb07253 Human nov
59	38	73.1	827	6	ABU07568	Abu07568 Human sec
60	38	73.1	904	4	ABG09947	Abg09947 Novel hum
61	38	73.1	924	5	AAE71323	Aae71323 Human GCR
62	38	73.1	953	7	ADE34415	Ade34415 Human G-p
63	38	73.1	994	5	ABB07252	Abb07252 Human nov
64	38	73.1	994	5	AAU99808	Aau99808 Novel hum
65	38	73.1	994	7	ADE34425	Ade34425 Human G-p
66	38	73.1	994	8	ADO28977	Ado28977 Human nov
67	38	73.1	994	8	ADQ25892	Adq25892 Human gua
68	38	73.1	1018	5	AAE25061	Aae25061 Human G-p
69	38	73.1	1070	6	ABU07567	Abu07567 Human sec
70	38	73.1	1131	4	ABG11655	Abg11655 Novel hum
71	38	73.1	1232	7	ADF70474	Adf70474 Orphan re
72	37	71.2	9	9	ADM12595	Adm12595 M32-40/DE
73	37	71.2	21	9	ADM12594	Adm12594 M20-40/DE
74	37	71.2	32	9	ADM12593	Adm12593 M10-40/DE
75	37	71.2	39	9	ADM12576	Adm12576 M1-40/DEN
76	37	71.2	40	5	AAE17432	Aae17432 Dengue (D
77	37	71.2	40	9	ADM12578	Adm12578 Dengue (D
78	37	71.2	48	5	AAE17433	Aae17433 (95-114)E
79	37	71.2	75	8	ADM97136	Adm97136 Japanese-
80	37	71.2	167	8	ADM97135	Adm97135 Japanese-
81	37	71.2	167	8	ADN37497	Adn37497 Dengue vi
82	37	71.2	171	8	ADN37493	Adn37493 Dengue vi
83	37	71.2	171	8	ADN37496	Adn37496 Dengue vi
84	37	71.2	319	8	ADR43186	Adr43186 IPT-like
85	37	71.2	319	10	AEF15713	Aef15713 Isopenthae
86	37	71.2	482	7	ADC01090	Adc01090 Enterohaer
87	37	71.2	484	3	AAW79292	Aaw79292 Mung bean
88	37	71.2	635	2	AAW75410	Aaw75410 Fusion pr
89	37	71.2	675	8	ADN37628	Adn37628 Dengue vi
90	37	71.2	675	8	ADN37518	Adn37518 Dengue vi
91	37	71.2	675	8	ADN37612	Adn37612 Dengue vi
92	37	71.2	675	8	ADN37626	Adn37626 Dengue vi
93	37	71.2	677	2	AAW75411	Aaw75411 Fusion pr
94	37	71.2	677	8	ADN37613	Adn37613 Dengue vi
95	37	71.2	681	8	ADN37603	Adn37603 Dengue vi
96	37	71.2	681	8	ADN37517	Adn37517 Dengue vi

97 37 71.2 682 8 ADW23990 Japanese  
 98 37 71.2 685 6 ABP57874 Plasmid p  
 99 37 71.2 685 6 ABP57876 Plasmid p  
 100 37 71.2 685 6 ABP57875

## ALIGNMENTS

RESULT 1  
 ADN37495  
 ID ADN37495 standard; protein; 167 AA.  
 XX  
 AC ADN37495;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Dengue virus type 4 (DEN-4) C15/truncated prM antigen fusion protein.  
 XX  
 KW virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;  
 KW capsid; DEN-4.  
 XX  
 OS Dengue virus type 4.  
 XX  
 PN WO2003102166-A2.  
 XX  
 PD 11-DEC-2003.  
 XX  
 PF 26-FEB-2003; 2003WO-US005918.  
 XX  
 PR 26-FEB-2002; 2002US-0360030P.  
 XX  
 PA (MAXY-) MAXYGEN INC.  
 XX  
 PI Apt D, Punnonen J, Brinkman AM;  
 XX  
 DR WPI; 2004-043106/04.  
 XX  
 XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX  
 PS Disclosure; SEQ ID NO 120; 409pp; English.  
 XX  
 CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of a Dengue virus type 4 (DEN-4) C15/truncated prM  
 CC antigen fusion protein of the invention which comprises the C-terminal 15  
 CC amino acids of the capsid protein fused to a truncated form of the prM  
 CC protein lacking the C-terminal 15 amino acids.  
 XX  
 SQ Sequence 167 AA;

Query Match 88.5%; Score 46; DB 8; Length 167;  
 Best Local Similarity 88.9%; Pred. No. 2.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
 |||||  
 Db 139 VESWILRNP 147

RESULT 2  
 ADN37499  
 ID ADN37499 standard; protein; 167 AA.  
 XX  
 AC ADN37499;

XX 17-JUN-2004 (first entry)  
 DT Dengue virus C15/truncated prM antigen fusion protein - SEQ ID 124.  
 DE virucide; Flavivirus; arboviruses group B; gene therapy; truncated prM;  
 KW capsid.  
 XX  
 OS Dengue virus.  
 XX  
 PN WO2003102166-A2.  
 XX  
 PD 11-DEC-2003.  
 XX  
 PF 26-FEB-2003; 2003WO-US005918.  
 XX  
 PR 26-FEB-2002; 2002US-0360030P.  
 XX  
 PA (MAXY-) MAXYGEN INC.  
 XX  
 PI Apt D, Punnonen J, Brinkman AM;  
 XX  
 DR WPI; 2004-043106/04.  
 XX  
 XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.  
 XX  
 PS Disclosure; SEQ ID NO 124; 409pp; English.  
 XX  
 CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of a Dengue virus C15/truncated prM antigen fusion  
 CC protein of the invention which comprises the C-terminal 15 amino acids of  
 CC the capsid protein fused to a truncated form of the prM protein lacking  
 CC the C-terminal 15 amino acids.  
 XX  
 SQ Sequence 167 AA;

Query Match 88.5%; Score 46; DB 8; Length 167;  
 Best Local Similarity 88.9%; Pred. No. 2.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
 |||||  
 Db 139 VESWILRNP 147

RESULT 3  
 AAW75412  
 ID AAW75412 standard; peptide; 635 AA.  
 XX  
 AC AAW75412;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 02-MAR-1999 (first entry)  
 XX  
 DE Fusion protein PD33 contains Dengue virus epitope.  
 XX  
 KW Dengue virus; fusion protein; P64K; Neisseria meningitidis; epitope;  
 KW antibody; diagnosis; Flavivirus; infection; vaccine.  
 XX  
 OS Dengue virus.  
 OS Neisseria meningitidis.  
 OS Chimeric.  
 XX

```

PN WO9831014-A1.
XX
XX 23-JUL-1998.
XX
XX 13-JAN-1998; 98WO-CU000001.
XX
XX 15-JAN-1997; 97CU-00000013.
XX
XX (CIGB-) CIGB CENT ING GENETICA & BIOTECNOLOGIA.
PA (IPKM-) IPK INST MEDICINA TROPICAL KOURI PEDRO.
XX
XX Vazquez Ramudo S, Guzman Tirado G, Guillen Nieto GE, Pardo Lazo OL;
PI Chinae Santiago G, Perez Diaz AB, Pupo Antunez M, Rodriguez Roche R;
PI Reyes Acosta O, Garay Perez HE, Padron Palomares G, Alvarez Vera M;
PI Morier Diaz L, Perez Insuaita O, Pelegrino Martinez De La Coterri Pedro;
XX
XX WPI; 1998-414111/35.
XX
XX New peptide(s) and fusion proteins useful for diagnosis and treatment of
PT flavivirus infection - contain cross-reactive epitopes from Dengue virus
PT pre-M/M protein and can induce neutralising antibodies.
XX
XX Claim 7; Page 32-34; 64pp; Spanish.
XX
XX This protein represents a fusion protein comprising 2 M protein epitopes
CC from Dengue virus type 2 and type 4 inserted into the P64K protein from
CC Neisseria meningitidis. Synthetic peptides based on the Dengue virus
CC epitope sequences (AAW75404-W75408) and fusion proteins can be used to
CC raise antibodies. The peptides, protein and antibodies are all useful for
CC diagnosis and treatment of Flavivirus infection, e.g. in vaccines.
CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 635 AA;

Query Match 88.5%; Score 46; DB 2; Length 635;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db ||||| |||||
74 VESWILRNP 82

RESULT 4
ADN37625
ID ADN37625 standard; protein; 675 AA.
XX
XX ADN37625;
XX
XX 17-JUN-2004 (first entry)
XX
XX DEN-4/Dengue virus C15/prM/E part codon-opt antigen fusion protein 1.
XX
XX virucide; Flavivirus; arboviruses group B; gene therapy; DEN-4;
XX C15/prM/E; human codon-optimised; prM; envelope; capsid.
XX
XX Dengue virus type 4.
XX
XX Dengue virus.
XX
XX Synthetic.
XX
XX WO2003102166-A2.
XX
XX 11-DEC-2003.
XX
XX 26-FEB-2003; 2003WO-US005918.
XX
XX 26-FEB-2002; 2002US-0360030P.
XX
XX (MAXY-) MAXYGEN INC.
XX
XX Apt D, Punnonen J, Brinkman AM;
XX
XX WPI; 2004-043106/04.
XX
XX N-PSDB; ADN37631.
XX
XX New recombinant or synthetic polypeptides and polynucleotides useful for
PT diagnosing, preventing or treating diseases associated with flaviviruses,
XX including dengue viruses.
XX
XX Example 28; SEQ ID NO 252; 409pp; English.
XX
XX

```

---

```

DR WPI; 2004-043106/04.
XX N-PSDB; ADN37629.
XX
XX New recombinant or synthetic polypeptides and polynucleotides useful for
PT diagnosing, preventing or treating diseases associated with flaviviruses,
XX including dengue viruses.
XX
XX Example 28; SEQ ID NO 250; 409pp; English.
XX
XX The invention relates to a novel recombinant or synthetic polypeptide
CC comprising an amino acid sequence that has at least about 90% sequence
CC identity to any of the 20 fully defined amino acid sequences given in the
CC specification. The polypeptide of the invention demonstrates virucide
CC activity and may be useful for inducing an immune response to
CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
CC in detecting and/or diagnosing the presence of antibodies against the
CC Dengue virus serotypes in a sample and for gene therapy. The current
CC sequence is that of the Dengue virus type 1 (DEN-4)/unidentified Dengue
CC virus C15/prM/E partially human codon-optimised antigen fusion protein of
CC the invention which comprises 15 amino acids of the DEN-4 capsid (C)
CC protein fused to the full-length DEN-4 codon-optimised/unidentified
CC Dengue virus prM protein and DEN-4 codon-optimised/unidentified Dengue
CC virus envelope (E) protein.
XX
XX Sequence 675 AA;

Query Match 88.5%; Score 46; DB 8; Length 675;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db ||||| |||||
139 VESWILRNP 147

RESULT 5
ADN37627
ID ADN37627 standard; protein; 675 AA.
XX
XX ADN37627;
XX
XX 17-JUN-2004 (first entry)
XX
XX DEN-4/Dengue virus C15/prM/E part codon-opt antigen fusion protein 2.
XX
XX virucide; Flavivirus; arboviruses group B; gene therapy; DEN-4;
XX C15/prM/E; human codon-optimised; prM; envelope; capsid.
XX
XX Dengue virus type 4.
XX
XX Dengue virus.
XX
XX Synthetic.
XX
XX WO2003102166-A2.
XX
XX 11-DEC-2003.
XX
XX 26-FEB-2003; 2003WO-US005918.
XX
XX 26-FEB-2002; 2002US-0360030P.
XX
XX (MAXY-) MAXYGEN INC.
XX
XX Apt D, Punnonen J, Brinkman AM;
XX
XX WPI; 2004-043106/04.
XX
XX N-PSDB; ADN37631.
XX
XX New recombinant or synthetic polypeptides and polynucleotides useful for
PT diagnosing, preventing or treating diseases associated with flaviviruses,
XX including dengue viruses.
XX
XX Example 28; SEQ ID NO 252; 409pp; English.
XX
XX

```

CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus type 1 (DEN-4)/unidentified Dengue  
 CC virus C15/prm/E partially human codon-optimised antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the DEN-4 capsid (C)  
 CC protein fused to the full-length DEN-4 codon-optimised/unidentified  
 CC Dengue virus prM protein and DEN-4 codon-optimised/unidentified Dengue  
 CC virus envelope (E) protein.

XX SQ Sequence 675 AA;

Query Match 88.5%; Score 46; DB 8; Length 675;  
 Best Local Similarity 88.9%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VESWFLRNP 9  
 |||||  
 Db 139 VESWILRNP 147

RESULT 6  
 ADN37514  
 ID ADN37514 standard; protein; 675 AA.

XX AC ADN37514;

DT 17-JUN-2004 (first entry)

XX Dengue virus C15/prm/E antigen fusion protein - SEQ ID 139.

DE virucide; Flavivirus; arboviruses group B; gene therapy; C15/prm/E; prM;  
 KW envelope; capsid.

XX OS Dengue virus.

XX PN WO2003102166-A2.

XX PD 11-DEC-2003.

XX PF 26-FEB-2003; 2003WO-US0005918.

XX PR 26-FEB-2002; 2002US-0360030P.

XX PA (MAXY-) MAXYGEN INC.

XX PI Apt D, Punnonen J, Brinkman AM;

XX DR WPI; 2004-043106/04.

XX DR N-PSDB; ADN37579.

XX PT New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

XX PS Claim 40; SEQ ID NO 139; 409pp; English.

XX CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prm/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.

XX SQ Sequence 675 AA;

Query Match 88.5%; Score 46; DB 8; Length 675;  
 Best Local Similarity 88.9%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VESWFLRNP 9  
 |||||  
 Db 139 VESWILRNP 147

RESULT 7

ADN37520  
 ID ADN37520 standard; protein; 675 AA.

XX AC ADN37520;

DT 17-JUN-2004 (first entry)

XX Dengue virus C15/prm/E antigen fusion protein - SEQ ID 145.

DE virucide; Flavivirus; arboviruses group B; gene therapy; C15/prm/E; prM;  
 KW envelope; capsid.

XX OS Dengue virus.

XX PN WO2003102166-A2.

XX PD 11-DEC-2003.

XX PF 26-FEB-2003; 2003WO-US0005918.

XX PR 26-FEB-2002; 2002US-0360030P.

XX PA (MAXY-) MAXYGEN INC.

XX PI Apt D, Punnonen J, Brinkman AM;

XX DR WPI; 2004-043106/04.

XX PT New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

XX PS Claim 40; SEQ ID NO 145; 409pp; English.

XX CC The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prm/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.

XX SQ Sequence 675 AA;

Query Match 88.5%; Score 46; DB 8; Length 675;  
 Best Local Similarity 88.9%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VESWFLRNP 9  
 |||||  
 Db 139 VESWILRNP 147

RESULT 8

ADN37516  
 ID ADN37516 standard; protein; 676 AA.

```

XX AC ADN37516;
XX DT 17-JUN-2004 (first entry)
XX DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 141.
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;
XX KW envelope; capsid.
XX OS Dengue virus.
XX PN WO2003102166-A2.
XX PD 11-DEC-2003.
XX PF 26-FEB-2003; 2003WO-US005918.
XX PR 26-FEB-2002; 2002US-0360030P.
XX PA (MAXY-) MAXYGEN INC.
XX PI Apt D, Punnonen J, Brinkman AM;
XX WPI; 2004-043106/04.
XX DR N-PSDB; ADN37577.
XX DR New recombinant or synthetic polypeptides and polynucleotides useful for
XX PT diagnosing, preventing or treating diseases associated with flaviviruses,
XX PT including dengue viruses.
XX FS Claim 40; SEQ ID NO 141; 409pp; English.
XX CC The invention relates to a novel recombinant or synthetic polypeptide
XX CC comprising an amino acid sequence that has at least about 90% sequence
XX CC identity to any of the 20 fully defined amino acid sequences given in the
XX CC specification. The polypeptide of the invention demonstrates virucide
XX CC activity and may be useful for inducing an immune response to
XX CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as
XX CC in detecting and/or diagnosing the presence of antibodies against the
XX CC Dengue virus serotypes in a sample and for gene therapy. The current
XX CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of
XX CC the invention which comprises 15 amino acids of the capsid (C) protein
XX CC fused to the full-length prM protein and envelope (E) protein.
XX SQ Sequence 676 AA;

Query Match 88.5%; Score 45; DB 8; Length 676;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 139 VESWILRNP 147

RESULT 9
ADN37605
ID ADN37605 standard; protein; 677 AA.
XX AC
XX AC
XX DT 17-JUN-2004 (first entry)
XX DE Dengue virus type 4 Den-4C15/prM/E antigen fusion protein.
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; DEN-4;
XX KW Den-4C15/prM/E; prM; envelope; capsid.
XX OS Dengue virus type 4.
XX PN WO2003102166-A2.
XX PT New recombinant or synthetic polypeptides and polynucleotides useful for

Query Match 88.5%; Score 46; DB 8; Length 677;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 139 VESWILRNP 147

RESULT 10
ADN37718
ID ADN37718 standard; protein; 677 AA.
XX AC
XX AC
XX DT 17-JUN-2004 (first entry)
XX DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 343.
XX KW virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;
XX KW envelope; capsid.
XX OS Dengue virus.
XX PN WO2003102166-A2.
XX PD 11-DEC-2003.
XX PF 26-FEB-2003; 2003WO-US005918.
XX PR 26-FEB-2002; 2002US-0360030P.
XX PA (MAXY-) MAXYGEN INC.
XX PI Apt D, Punnonen J, Brinkman AM;
XX WPI; 2004-043106/04.
XX DR N-PSDB; ADN37717.
XX PT New recombinant or synthetic polypeptides and polynucleotides useful for

```

PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

PS Claim 40; SEQ ID NO 343; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.

XX Sequence 677 AA;

Query Match 88.5%; Score 46; DB 8; Length 677;  
 Best Local Similarity 88.9%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9  
 |||||  
 Db 139 VESWILRNP 147

RESULT 11

ADN37720  
 ID ADN37720 standard; protein; 677 AA.

XX

AC ADN37720;

DT 17-JUN-2004 (first entry)

DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 345.

XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;  
 KW envelope; capsid.

XX Dengue virus.

XX WO2003102166-A2.

XX 11-DEC-2003.

XX 26-FEB-2003; 2003WO-US005918.

XX 26-FEB-2002; 2002US-0360030P.

XX (MAXY-) MAXYGEN INC.

XX Apt D, Punnonen J, Brinkman AM;

XX WPI; 2004-043106/04.

DR N-PSDB; ADN37719.

XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

XX Claim 40; SEQ ID NO 345; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of

CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.

XX Sequence 677 AA;

Query Match 88.5%; Score 46; DB 8; Length 677;  
 Best Local Similarity 88.9%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9  
 |||||  
 Db 139 VESWILRNP 147

RESULT 12

ADN37623

ID ADN37623 standard; protein; 678 AA.

XX

AC ADN37623;

DT 17-JUN-2004 (first entry)

DE Dengue virus C15/prM/E antigen fusion protein - SEQ ID 248.

XX virucide; Flavivirus; arboviruses group B; gene therapy; C15/prM/E; prM;  
 KW envelope; capsid.

XX Dengue virus.

XX WO2003102166-A2.

XX 11-DEC-2003.

XX 26-FEB-2003; 2003WO-US005918.

XX 26-FEB-2002; 2002US-0360030P.

XX (MAXY-) MAXYGEN INC.

XX Apt D, Punnonen J, Brinkman AM;

XX WPI; 2004-043106/04.

XX New recombinant or synthetic polypeptides and polynucleotides useful for  
 PT diagnosing, preventing or treating diseases associated with flaviviruses,  
 PT including dengue viruses.

XX Example 13; SEQ ID NO 248; 409pp; English.

XX The invention relates to a novel recombinant or synthetic polypeptide  
 CC comprising an amino acid sequence that has at least about 90% sequence  
 CC identity to any of the 20 fully defined amino acid sequences given in the  
 CC specification. The polypeptide of the invention demonstrates virucide  
 CC activity and may be useful for inducing an immune response to  
 CC Flaviviruses (arboviruses group B), including Dengue viruses, as well as  
 CC in detecting and/or diagnosing the presence of antibodies against the  
 CC Dengue virus serotypes in a sample and for gene therapy. The current  
 CC sequence is that of the Dengue virus C15/prM/E antigen fusion protein of  
 CC the invention which comprises 15 amino acids of the capsid (C) protein  
 CC fused to the full-length prM protein and envelope (E) protein.

XX Sequence 678 AA;

Query Match 88.5%; Score 46; DB 8; Length 678;  
 Best Local Similarity 88.9%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9  
 |||||  
 Db 142 VESWILRNP 150

RESULT 13



AAE07992  
ID AAE07992 standard; protein; 3387 AA.  
XX  
AC AAE07992;  
XX  
DT 01-NOV-2001 (first entry)  
XX  
DE Attenuated, vaccine-strain DEN-4 PDK-48 protein variant.  
XX  
DE Flavivirus; Dengue virus-4; DEN-4; vaccine; infection; virucidal; muten;  
KW avirulent; immunogenic; viral disease; pharmaceutical; mutant; variant.  
XX  
OS Dengue virus; type IV.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 624  
FT /note= "Wild type Glu substituted with Lys"  
FT Misc-difference 1027  
FT /note= "Wild type Gln substituted with His"  
FT Misc-difference 2187  
FT /note= "Wild type Leu substituted with Phe"  
FT Misc-difference 2286  
FT /note= "Wild type Ile substituted with Phe"  
FT Misc-difference 2354  
FT /note= "Wild type Leu substituted with Ser"  
FT Misc-difference 2366  
FT /note= "Wild type Ala substituted with Val"  
FT Misc-difference 2508  
FT /note= "Wild type Asp substituted with Tyr"  
XX  
PN WO200160847-A2.  
XX  
PD 23-AUG-2001.  
XX  
XX 16-FEB-2001; 2001WO-US005142.  
XX  
XX 16-FEB-2000; 2000US-0182829P.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Kinney RM, Kinney CYH, Butrapet S, Gubler DL, Bhamarapravati N;  
XX  
XX WPI: 2001-497162/54.  
XX N-PSDB; AAD14613.  
XX  
XX Chimeric flaviviruses that are avirulent and immunogenic, useful for  
PT vaccinating against a range of dengue viruses.  
XX  
XX Example 4; Page 414-421; 470pp; English.  
XX  
XX The invention relates to avirulent, immunogenic flavivirus chimeras  
CC comprising amino acid mutations in the non-structural proteins of a  
CC flavivirus. Chimeric viruses containing the attenuation-mutated non-  
CC structural genes of the virus are used as a backbone into which the  
CC structural protein genes of a second flavivirus strain are inserted.  
CC These chimeric viruses elicit pronounced immunogenicity but lack the  
CC accompanying clinical symptoms of viral disease. Attenuated chimeric  
CC flaviviruses are combined in a pharmaceutical composition to confer  
CC simultaneous immunity against several strains of pathogenic flaviviruses  
CC such as dengue virus serotypes DEN-1, DEN-2, DEN-3 and DEN-4. Immunogenic  
CC flavivirus chimeras are also used as immunogens or multivalent vaccines  
CC to confer simultaneous protection against infections. The present  
CC sequence is attenuated dengue-4 (DEN-4) PDK-48 virus protein variant used  
CC for constructing flavivirus chimeras. Dengue virus types 1-4 (DEN-1 to  
CC DEN-4) are mosquito-borne flavivirus pathogens. The flavivirus protein  
CC contains a capsid protein (C), premembrane/membrane protein (prM), an  
CC envelope protein (E) and non-structural proteins (NS1-NS2A-NS2B-NS3-NS4A-  
CC NS4B-NS5). DEN-4 virus passaged in primary dog kidney (PDK) cells 48  
XX times is designated as DEN-4 PDK-48 virus  
SQ Sequence 3387 AA;

Query Match 88.5%; Score 46; DB 4; Length 3387;  
Best Local Similarity 88.9%; Pred. No. 76;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VESWFLRNP 9  
Db 236 VESWILRNP 244  
RESULT 16  
AAE35313  
ID AAE35313 standard; protein; 3387 AA.  
XX  
AC AAE35313;  
XX  
DT 28-MAY-2003 (first entry)  
XX  
DE Recombinant dengue virus type 4 strain rDEN4 protein.  
XX  
KW Attenuation; growth; vaccine; infection; Dengue virus type 4.  
XX  
OS Dengue virus.  
XX  
PN WO200295075-A1.  
XX  
PD 28-NOV-2002.  
XX  
PF 22-MAY-2002; 2002WO-US016308.  
XX  
PR 22-MAY-2001; 2001US-0293049P.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX (BLAN/) BLANEY J E.  
XX  
XX Whitehead SS, Murphy BR, Hanley KA;  
XX  
XX WPI: 2003-120809/11.  
XX N-PSDB; AAD53911.  
XX  
XX New mutated flavivirus, useful for fine tuning the attenuation and growth  
PT characteristics of dengue virus vaccines for the prevention and/or  
PT treatment of dengue virus infection.  
XX  
XX Disclosure; Page 127-128; 246pp; English.  
XX  
XX The present invention relates to novel mutated flaviviruses comprising a  
CC phenotype in which the viral genome is modified by introduction of a  
CC mutation, singly or in combination, taken from mutations from recombinant  
CC virus bearing Vero adaptation mutations, putative Vero cell adaptation  
CC mutations of dengue type 4 virus (DEN4) or mutations known to attenuate  
CC dengue type 4 virus. The methods and compositions of the invention are  
CC useful for fine tuning the attenuation and growth characteristics of  
CC dengue virus vaccines for the prevention and/or treatment of dengue virus  
CC infection. The present sequence is Dengue virus type 4 strain rDEN4  
XX protein  
XX  
XX Sequence 3387 AA;  
SQ

Query Match 88.5%; Score 46; DB 6; Length 3387;  
Best Local Similarity 88.9%; Pred. No. 76;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VESWFLRNP 9  
Db 236 VESWILRNP 244  
RESULT 17  
AAE35312  
ID AAE35312 standard; protein; 3387 AA.  
XX  
AC AAE35312;  
XX





```
DR WPI; 2005-047647/05.
XX
PT New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
PT as a vaccine for preventing or treating pathological conditions from non-
PT specific febrile illnesses to severe hemorrhagic manifestations or
PT encephalitic syndromes.
XX
XX Example 1; SEQ ID NO 29; 30pp; English.
XX
CC The present invention relates to an isolated and purified ApoptoM
CC peptide. The invention is useful as a vaccine for the prevention and
CC treatment of pathological conditions from non-specific febrile illnesses
CC to severe hemorrhagic manifestations, encephalitic syndromes and these
CC pathological conditions are linked to Flavivirus infection or cancers.
CC The invention is also useful in gene therapy. The present sequence is a
CC M1-40/DEN (dengue)-2 (F36) mutant protein.
XX
XX Sequence 39 AA;
SQ
Query Match 82.7%; Score 43; DB 9; Length 39;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 VESWFLRNP 9
Db 31 IETWFLRHP 39
:|:|:|:|
RESULT 20
ADW12577
ID ADW12577 standard; peptide; 40 AA.
XX
AC ADW12577;
XX
DT 24-MAR-2005 (first entry)
XX
DE M1-40/YF.17D protein.
XX
XX Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX YF; yellow fever.
XX
XX Yellow fever virus.
XX
XX US2004266987-A1.
XX
XX 30-DEC-2004.
XX
XX 30-JUN-2003; 2003US-00608029.
XX
XX 30-JUN-2003; 2003US-00608029.
XX
XX (INSP ) INST PASTEUR.
XX
XX Despres P, Catteau A;
XX
XX WPI; 2005-047647/05.
XX
XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
XX as a vaccine for preventing or treating pathological conditions from non-
XX specific febrile illnesses to severe hemorrhagic manifestations or
XX encephalitic syndromes.
XX
XX Example 3; SEQ ID NO 24; 30pp; English.
XX
CC The present invention relates to an isolated and purified ApoptoM
CC peptide. The invention is useful as a vaccine for the prevention and
CC treatment of pathological conditions from non-specific febrile illnesses
CC to severe hemorrhagic manifestations, encephalitic syndromes and these
CC pathological conditions are linked to Flavivirus infection or cancers.
CC The invention is also useful in gene therapy. The present sequence is a
CC M1-40/YF (yellow fever).17D protein.
XX
XX
```

```
SQ Sequence 40 AA;
Query Match 82.7%; Score 43; DB 9; Length 40;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VESWFLRNP 9
Db 32 IERWFLRNP 40
:|:|:|:|
RESULT 21
ADW12588
ID ADW12588 standard; protein; 48 AA.
XX
AC ADW12588;
XX
DT 24-MAR-2005 (first entry)
XX
DE p(95-114) EGFP (M1-M40) DEN-2 (136F) plasmid DNA encoded protein #3.
XX
XX Gene therapy; protein purification; virucide; cytostatic; vaccine;
XX hemorrhage; bleeding; infection; flavivirus infection; neoplasm; cancer;
XX DEN; dengue; EGFP; enhanced green fluorescent protein.
XX
XX Dengue virus.
XX
XX Chimeric.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FT Misc-difference 2 /note= "Encoded by GGC"
FT Misc-difference 4 /note= "Encoded by GAC"
FT Misc-difference 13.44 /note= "Encoded by GTTTC"
FT
FT
XX
XX US2004266987-A1.
XX
XX 30-DEC-2004.
XX
XX 30-JUN-2003; 2003US-00608029.
XX
XX 30-JUN-2003; 2003US-00608029.
XX
XX (INSP ) INST PASTEUR.
XX
XX Despres P, Catteau A;
XX
XX WPI; 2005-047647/05.
XX
XX N-PSDB; ADW12589.
XX
XX New isolated and purified ApoptoM peptide comprises 9 amino acids, useful
XX as a vaccine for preventing or treating pathological conditions from non-
XX specific febrile illnesses to severe hemorrhagic manifestations or
XX encephalitic syndromes.
XX
XX Disclosure; SEQ ID NO 35; 30pp; English.
XX
XX The present invention relates to an isolated and purified ApoptoM
XX peptide. The invention is useful as a vaccine for the prevention and
XX treatment of pathological conditions from non-specific febrile illnesses
XX to severe hemorrhagic manifestations, encephalitic syndromes and these
XX pathological conditions are linked to Flavivirus infection or cancers.
XX The invention is also useful in gene therapy. The present sequence is a
XX p(95-114) EGFP (enhanced green fluorescent protein) (M1-M40) DEN (dengue)-2
XX (136F) plasmid DNA encoded protein.
XX
XX Sequence 48 AA;
SQ
Query Match 82.7%; Score 43; DB 9; Length 48;
Best Local Similarity 66.7%; Pred. No. 2.4;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 VESWFLRNP 9
   :|||:|
DB 40 IETWFLRHP 48

RESULT 22
ADR87182
ID ADR87182 standard; protein; 680 AA.
XX
AC ADR87182;
XX
DT 18-NOV-2004 (first entry)
XX
DE Yellow Fever Virus deltaCprME, SEQ ID 9.
XX
KW Virucide; Vaccine; Gene therapy; flavivirus; flavivirus core protein;
KW flavivirus prM protein; flavivirus E protein; envelope glycoprotein;
KW deltaCprME.
XX
OS Synthetic.
XX
PN EPI454988-A1.
XX
PD 08-SEP-2004.
XX
PF 03-MAR-2003; 2003EP-00290504.
XX
PR 03-MAR-2003; 2003EP-00290504.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Bartosch B, Cosset F;
XX
DR WPI; 2004-654727/64.
DR N-PSDB; ADR87181.
XX

Producing flavivirus-like particles for use as a vaccine, comprises
transfecting cells with cDNA encoding flavivirus prM and/or E protein,
PT and expressing and allowing structural proteins to form virus-like
PT particles.
XX
Example 1; SEQ ID NO 9; 38pp; English.
XX
The present invention relates to a method for producing flavivirus-like
particles (I) ex vivo, by providing nucleic acids comprising packaging
competent retroviral-derived genome, cDNA encoding core proteins from
retrovirus, and cDNA encoding a polyprotein comprising flavivirus core
protein and flavivirus prM protein and/or a flavivirus E envelope
glycoproteins (ADR87177-ADR87182). The packaging competent retroviral-
derived genome and core proteins are form a retrovirus chosen from murine
leukaemia virus (MLV), ALV, respiratory syncytial virus (RSV), MPV, HIV-
1, HIV-2, SIV, EIAV, CAEV, or HFV. The flavivirus core protein is a
carboxy-terminus of flavivirus core that comprises the core protein
signal peptide (ADR87174-ADR87176). The polyprotein comprises a
flavivirus core protein and a native flavivirus prM and/or E protein. The
core, prM and E flavivirus proteins are derived from a same flavivirus.
The flavivirus is chosen from West Nile virus, Dengue virus and yellow
fever virus. The nucleic acid sequence comprising a packaging competent
retroviral-derived genome further comprises a transgene. (I) are useful
for ex vivo identification of a receptor for flavivirus prM and/or E
glycoprotein; for ex vivo screening or identification of molecules
capable of interfering with flavivirus entry in cells; for in vitro
diagnosis of a flavivirus infection in a patient; for in vitro
transferring of a transgene of interest in a target cell and as a vector
for gene transfer and/or gene therapy. Flavivirus pseudo-particles (FVpp)
were generated by assembling full-length, unmodified prM and E
glycoproteins onto retroviral core proteins derived from murine leukaemia
virus (MLV). The pCMV-deltaC prME-WNV expression vector encoding the prM
and E glycoproteins from West Nile Virus was generated by inserting a
blunt-ended fragment encoding the last 22 residues of flavivirus core (C)
and all of prM and E proteins into the BamHI digested and Klenow blunted

```

```

CC vector pCMV-G. The present sequence is Yellow Fever Virus deltaCprME
CC protein comprising successively a carboxy terminus of flavivirus core
CC protein, and flavivirus prM and flavivirus E proteins.
XX
SQ Sequence 680 AA;

Query Match      82.7%; Score 43; DB 8; Length 680;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
   :|||:|
DB 144 IERWFLRNP 152

RESULT 23
ABP57863
ID ABP57863 standard; protein; 681 AA.
XX
AC ABP57863;
XX
DT 07-FEB-2003 (first entry)
XX
DE Plasmid pCBYF containing yellow fever virus prM and E.
XX
KW Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;
KW pCBYF; yellow fever virus.
XX
OS Unidentified.
OS Yellow fever virus.
OS Chimeric.
XX
PN WO200281754-A1.
XX
PD 17-OCT-2002.
XX
PF 04-APR-2002; 2002WO-US010764.
XX
PR 04-APR-2001; 2001US-00826115.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Chang GJ;
XX
DR WPI; 2003-058572/05.
DR N-PSDB; ABV77540.
XX
PT Novel isolated nucleic acid useful as vaccine for preventing flavivirus
PT infection, comprises transcriptional unit encoding signal sequence of one
PT flavivirus and immunogenic flavivirus antigen of a second flavivirus.
XX
Example 14; Page 147-148; 174pp; English.
XX
The invention relates to a novel nucleic acid comprising a
transcriptional unit encoding a signal sequence of a structural protein
of a first flavivirus and an immunogenic flavivirus antigen of a second
flavivirus, where the transcriptional unit directs the synthesis of the
antigen. The polynucleotide of the invention has virucide activity, and
acts as a vaccine. A composition of the invention is useful for
immunising a subject against infection by a flavivirus. The
polynucleotide is useful as a vaccine for preventing flavivirus
infection. The sequence represents plasmid pCBYF, which contains yellow
fever virus prM and E proteins
XX
SQ Sequence 681 AA;

Query Match      82.7%; Score 43; DB 6; Length 681;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
   :|||:|
DB 145 IERWFLRNP 153

```

```

RESULT 24
ABP57861
ID ABP57861 standard; protein; 681 AA.
XX AC ABP57861;
XX DT 07-FEB-2003 (first entry)
XX DE Plasmid pCBPOW containing Powassan virus prM and E.
XX KW Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;
XX KW pCBPOW; Powassan virus.
XX OS Unidentified.
XX OS Powassan virus.
XX OS Chimeric.
XX PN WO200281754-A1.
XX PD 17-OCT-2002.
XX PF 04-APR-2002; 2002WO-US010764.
XX PR 04-APR-2001; 2001US-00826115.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Chang GJ;
XX DR WPI; 2003-058572/05.
XX PT Novel isolated nucleic acid useful as vaccine for preventing flavivirus
PT infection, comprises transcriptional unit encoding signal sequence of one
PT flavivirus and immunogenic flavivirus antigen of a second flavivirus.
XX PS Example 15; Page 135-137; 174pp; English.
XX CC The invention relates to a novel nucleic acid comprising a
CC transcriptional unit encoding a signal sequence of a structural protein
CC of a first flavivirus and an immunogenic flavivirus antigen of a second
CC flavivirus, where the transcriptional unit directs the synthesis of the
CC antigen. The polynucleotide of the invention has virucide activity, and
CC acts as a vaccine. A composition of the invention is useful for
CC immunising a subject against infection by a flavivirus. The
CC polynucleotide is useful as a vaccine for preventing flavivirus
CC infection. The sequence represents plasmid pCBPOW, which contains
CC Powassan virus prM and E encoding proteins. Note: The protein sequence is
CC not encoded by the cDNA sequence given in ABV7538, which is quoted as
CC the encoding polynucleotide in the specification
XX SQ Sequence 681 AA;
Query Match 82.7%; Score 43; DB 6; Length 681;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VESWFLRNP 9
Db 145 IERWFRNP 153
: |||:|
: |||:|

RESULT 25
ADJ57394
ID ADJ57394 standard; protein; 3411 AA.
XX AC ADJ57394;
XX DT 06-MAY-2004 (first entry)
XX DE Hamster passage 7 yellow fever virus polyprotein.
XX

```

```

KW yellow fever; attenuation; vaccine.
XX
XX Yellow fever virus.
FH Key Location/Qualifiers
FT 286..778
FT /label= Envelope_protein
XX
PN WO2004009764-A2.
XX
XX 29-JAN-2004.
XX
XX 11-JUL-2003; 2003WO-US022001.
XX
XX 19-JUL-2002; 2002US-0397440P.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX Barrett A, McArthur M;
XX
XX WPI; 2004-132942/13.
XX N-PSDB; ADJ57393.
XX
XX New isolated nucleic acid encoding a Yellow Fever virus and having
XX alterations of at least two nucleotides, useful in the fields of
XX molecular biology and virology, particularly for producing an improved
XX Yellow Fever virus vaccine.
XX
XX Disclosure; SEQ ID NO 2; 109pp; English.
XX
XX The present sequence is that of an Asibi/hamster passage 7 yellow fever
XX virus polyprotein. Comparison of this sequence with that of an
XX Asibi/hamster passage 0 virus revealed 7 amino acid substitutions: in the
XX envelope protein, His replaces Gln-27, Gly replaces Asp-28, Ala replaces
XX Asp-155, Arg replaces Lys-323 and Arg replaces Lys-331; in the NS2A
XX protein, Ala replaces Thr-48; and in the NS4B protein, Ile replaces Val-
XX 98. A claimed isolated nucleic acid encodes a yellow fever virus with a
XX viral genome that comprises an altered nucleic acid sequence resulting in
XX at least 1, and optionally up to all 7, of these amino acid
XX substitutions. The virus is used in a vaccine composition. The methods
XX and compositions of the invention provide for improvement of the
XX reversion frequency of an attenuated yellow fever virus for use in safer
XX vaccines in which the risk of disease is reduced or eliminated.
XX
XX SQ Sequence 3411 AA;
Query Match 82.7%; Score 43; DB 8; Length 3411;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VESWFLRNP 9
Db 242 IERWFRNP 250
: |||:|
: |||:|

RESULT 26
ADW00905
ID ADW00905 standard; protein; 167 AA.
XX
XX AC ADW00905;
XX
XX 10-MAR-2005 (first entry)
XX
XX Amino acid sequence of a WNV prM region.
XX
XX WNV; prM protein; E protein; polyprotein; envelope; membrane; vaccine.
XX
XX West Nile virus.
XX
XX WO2004112694-A2.
XX
XX 29-DEC-2004.
XX

```

PF 21-MAY-2004; 2004WO-US015976.  
 XX  
 PR 23-MAY-2003; 2003US-0473225P.  
 PR 11-DEC-2003; 2003US-0529171P.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Andrews W, Chien DY, Choo Q, Coates SR, Coit D, Harrington C;  
 PI Hilt S, Houghton M, Medina-Selby A, Pichuanes S;  
 XX  
 DR WPI; 2005-075046/08.  
 DR N-PSDB; ADW00904.  
 XX  
 PT Isolated immunogenic composition useful for immunizing animal against  
 PT West Nile Virus WNV, comprises single or complex of WNV PrM/E heterodimer  
 PT consisting of recombinant WNV PrM polypeptide and recombinant WNV E  
 PT polypeptide.  
 XX  
 PS Disclosure; SEQ ID NO 9; 130pp; English.  
 XX  
 CC The specification describes an immunogenic composition which comprises a  
 CC complex of West Nile Virus (WNV) PrM/E heterodimer or a single WNV PrM/E  
 CC heterodimer consisting of a recombinant WNV PrM polypeptide and a  
 CC recombinant WNV E polypeptide. The immunogenic composition of the  
 CC invention is useful for immunizing an animal against the WNV. It is also  
 CC useful for detecting WNV antibodies in a biological sample. The present  
 CC sequence represents a WNV PrM region. The WNV polypeptide is  
 CC proteolytically processed by the viral serine protease NS2B-NS3 and  
 CC various cellular proteases into 10 mature viral proteins, in the order:  
 CC capsid (C)-membrane (PrM)-envelope (E)- nonstructural protein 1 (NS1)-  
 CC NS2A-NS2B-NS3-NS4A-NS4B-NS5.  
 XX  
 SQ Sequence 167 AA;  
 Query Match 80.8%; Score 42; DB 9; Length 167;  
 Best Local Similarity 87.5%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 ESWFLRNP 9  
 ||| ||||  
 Db 125 ESWILRNP 132

RESULT 27  
 ADW00902  
 ID ADW00902 standard; protein; 668 AA.  
 XX  
 AC ADW00902;  
 XX  
 DT 10-MAR-2005 (first entry)  
 XX  
 DE Amino acid sequence of synthetic WNV PrM/E region.  
 XX  
 KW WNV; PrM protein; E protein; polypeptide; envelope; membrane; vaccine.  
 XX  
 OS West Nile virus.  
 OS Synthetic.  
 XX  
 PN WO2004112694-A2.  
 XX  
 PD 29-DEC-2004.  
 XX  
 PF 21-MAY-2004; 2004WO-US015976.  
 XX  
 PR 23-MAY-2003; 2003US-0473225P.  
 PR 11-DEC-2003; 2003US-0529171P.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Andrews W, Chien DY, Choo Q, Coates SR, Coit D, Harrington C;  
 PI Hilt S, Houghton M, Medina-Selby A, Pichuanes S;  
 XX  
 DR WPI; 2005-075046/08.

DR N-PSDB; ADW00901.  
 XX  
 PT Isolated immunogenic composition useful for immunizing animal against  
 PT West Nile Virus WNV, comprises single or complex of WNV PrM/E heterodimer  
 PT consisting of recombinant WNV PrM polypeptide and recombinant WNV E  
 PT polypeptide.  
 XX  
 PS Example 2; SEQ ID NO 6; 130pp; English.  
 XX  
 CC The specification describes an immunogenic composition which comprises a  
 CC complex of West Nile Virus (WNV) PrM/E heterodimer or a single WNV PrM/E  
 CC heterodimer consisting of a recombinant WNV PrM polypeptide and a  
 CC recombinant WNV E polypeptide. The immunogenic composition of the  
 CC invention is useful for immunizing an animal against the WNV. It is also  
 CC useful for detecting WNV antibodies in a biological sample. The present  
 CC sequence represents a synthetic PrM/E region of WNV. The WNV polypeptide  
 CC is proteolytically processed by the viral serine protease NS2B-NS3 and  
 CC various cellular proteases into 10 mature viral proteins, in the order:  
 CC capsid (C)-membrane (PrM)-envelope (E)- nonstructural protein 1 (NS1)-  
 CC NS2A-NS2B-NS3-NS4A-NS4B-NS5.  
 XX  
 SQ Sequence 668 AA;  
 Query Match 80.8%; Score 42; DB 9; Length 668;  
 Best Local Similarity 87.5%; Pred. No. 67;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 ESWFLRNP 9  
 ||| ||||  
 Db 125 ESWILRNP 132

RESULT 28  
 ADS76177  
 ID ADS76177 standard; protein; 691 AA.  
 XX  
 AC ADS76177;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Heterodimer glycoprotein, PrM-E.  
 XX  
 KW secreted envelope protein; heterodimer glycoprotein; PrM-E;  
 KW NS-1 protein; West-Nile virus; measles virus; vaccine; Dengue virus;  
 KW antigen.  
 XX  
 OS West Nile virus.  
 XX  
 PN WO2004076619-A2.  
 XX  
 PD 10-SEP-2004.  
 XX  
 PF 26-FEB-2004; 2004WO-IB001027.  
 XX  
 PR 26-FEB-2003; 2003CA-02420092.  
 PR 20-JUN-2003; 2003CA-02432738.  
 XX  
 PA (INSP ) INST PASTEUR.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Tangy F, Despres F, Combredet C, Frankiel MP;  
 XX  
 DR WPI; 2004-653390/63.  
 DR N-PSDB; ADS76173.  
 XX  
 PT New purified polypeptide derived from a West-Nile or Dengue virus  
 PT antigen, useful in preparing a vaccine for diagnosing, treating and/or  
 PT preventing West-Nile or Dengue viral infections.  
 XX  
 PS Claim 5; SEQ ID NO 6; 64pp; English.  
 XX  
 CC This sequence represents the heterodimer glycoprotein, PrM-E, from West-  
 CC Nile virus. This protein was used as an antigen to raise a purified

CC polyclonal or monoclonal antibody. The cDNA encoding this protein may be  
 CC used in the generation of a measles virus which may be used in a  
 CC pharmaceutical composition, or in the production of a recombinant virus  
 CC for the preparation of an anti-West-Nile virus vaccine or an anti-Dengue  
 CC virus vaccine. The polynucleotide sequences of the invention are useful  
 CC for detecting the presence or absence of a West-Nile or Dengue virus  
 CC antigen in a biological sample. The pharmaceutical composition is an anti  
 CC -West-Nile or Dengue virus agent in the preparation of an anti-West-Nile  
 CC or Dengue virus vaccine. They can also be used in diagnosing and  
 CC preventing West-Nile or Dengue viral infections.

XX  
 SQ Sequence 691 AA;

Query Match 80.8%; Score 42; DB 8; Length 691;  
 Best Local Similarity 87.5%; Pred. No. 70;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ESWFLRNP 9  
 Db 148 ESWILRNP 155  
 ||| ||||

RESULT 29  
 ABP57859  
 ID ABP57859 standard; protein; 692 AA.

XX AC ABP57859;  
 XX DT 07-FEB-2003 (first entry)  
 XX DE Plasmid pCBWN containing West Nile virus prM and E.  
 XX KW Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;  
 XX KW pCBWN; West Nile virus.

XX OS Unidentified.  
 XX OS West Nile virus.  
 XX OS Chimeric.  
 XX PN WO200281754-A1.

XX PD 17-OCT-2002.  
 XX PF 04-APR-2002; 2002WO-US010764.  
 XX PR 04-APR-2001; 2001US-00826115.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Chang GJ;  
 XX DR WPI; 2003-058572/05.  
 XX DR N-PSDB; ABV77536.

XX PT Novel isolated nucleic acid useful as vaccine for preventing flavivirus  
 PT infection, comprises transcriptional unit encoding signal sequence of one  
 PT flavivirus and immunogenic flavivirus antigen of a second flavivirus.

XX PS Example 9; Page 124-125; 174pp; English.

XX CC The invention relates to a novel nucleic acid comprising a  
 CC transcriptional unit encoding a signal sequence of a structural protein  
 CC of a first flavivirus and an immunogenic flavivirus antigen of a second  
 CC flavivirus, where the transcriptional unit directs the synthesis of the  
 CC antigen. The polynucleotide of the invention has virucide activity, and  
 CC acts as a vaccine. A composition of the invention is useful for  
 CC immunising a subject against infection by a flavivirus. The  
 CC polynucleotide is useful as a vaccine for preventing flavivirus  
 CC infection. The sequence represents plasmid pCBWN, which contains West  
 CC Nile virus prM and E proteins

XX Sequence 692 AA;

Query Match 80.8%; Score 42; DB 6; Length 692;  
 Best Local Similarity 87.5%; Pred. No. 70;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ESWFLRNP 9  
 Db 149 ESWILRNP 156  
 ||| ||||

RESULT 30  
 ABP57862  
 ID ABP57862 standard; protein; 692 AA.

XX AC ABP57862;  
 XX DT 07-FEB-2003 (first entry)  
 XX DE Plasmid pCBJESS-M containing St. Louis encephalitis virus prM and E.  
 XX KW Flavivirus; immunogenic flavivirus antigen; virucide; vaccine; prM-E;  
 XX KW pCBJESS-M; St. Louis encephalitis virus.

XX OS Unidentified.  
 XX OS St. Louis encephalitis virus.  
 XX OS Chimeric.

XX PN WO200281754-A1.

XX PD 17-OCT-2002.

XX PF 04-APR-2002; 2002WO-US010764.

XX PR 04-APR-2001; 2001US-00826115.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Chang GJ;

XX DR WPI; 2003-058572/05.  
 XX DR N-PSDB; ABV77539.

XX PT Novel isolated nucleic acid useful as vaccine for preventing flavivirus  
 PT infection, comprises transcriptional unit encoding signal sequence of one  
 PT flavivirus and immunogenic flavivirus antigen of a second flavivirus.

XX PS Example 13; Page 141-142; 174pp; English.

XX CC The invention relates to a novel nucleic acid comprising a  
 CC transcriptional unit encoding a signal sequence of a structural protein  
 CC of a first flavivirus and an immunogenic flavivirus antigen of a second  
 CC flavivirus, where the transcriptional unit directs the synthesis of the  
 CC antigen. The polynucleotide of the invention has virucide activity, and  
 CC acts as a vaccine. A composition of the invention is useful for  
 CC immunising a subject against infection by a flavivirus. The  
 CC polynucleotide is useful as a vaccine for preventing flavivirus  
 CC infection. The sequence represents plasmid pCBJESS-M, which contains St.  
 CC Louis encephalitis virus prM and E proteins

XX SQ Sequence 692 AA;

Query Match 80.8%; Score 42; DB 6; Length 692;  
 Best Local Similarity 77.8%; Pred. No. 70;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9  
 Db 148 VENWVLRNP 156  
 ||: ||||

Search completed: August 31, 2006, 11:50:36  
 Job time : 112.25 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 11:43:31 ; Search time 17.25 Seconds  
(without alignments)  
50.200 Million cell updates/sec

Title: DENGUE\_SEROTYPE4

Perfect score: 52

Sequence: 1 veswflmp 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 80.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	88.5	773	2 A47666	structural polypro
2	46	88.5	3386	1 GNVWDF	genome polyprotein
3	44	84.6	1525	1 GNVWS5	genome polyprotein
4	43	82.7	3411	1 GNVWYI	genome polyprotein
5	43	82.7	3411	1 GNVWYP	genome polyprotein
6	42	80.8	3430	1 GNVWVW	genome polyprotein
7	42	80.8	3433	1 GNVWVW	genome polyprotein
8	40	76.9	73	2 D69104	conserved hypother
9	39	75.0	227	2 H69453	hypothetical prote
10	39	75.0	232	2 S58857	botulinum neurotox
11	39	75.0	582	2 F84807	hypothetical prote
12	37	71.2	94	2 T29563	hypothetical prote
13	37	71.2	166	2 S40144	premembrane protei
14	37	71.2	266	2 S67182	hypothetical prote
15	37	71.2	489	2 T06024	1-aminocyclopropan
16	37	71.2	470	2 S71174	1-aminocyclopropan
17	37	71.2	482	2 H91109	hypothetical prote
18	37	71.2	482	2 H85955	unknown protein en
19	37	71.2	484	2 S25002	1-aminocyclopropan
20	37	71.2	484	2 S26214	1-aminocyclopropan
21	37	71.2	555	2 JQ1404	genome polyprotein
22	37	71.2	775	2 A48644	polyprotein - deng
23	37	71.2	789	2 I50804	polyprotein - Japa
24	37	71.2	1163	1 GNVWY8	genome polyprotein
25	37	71.2	3388	1 GNVWDP	genome polyprotein
26	37	71.2	3391	1 GNVWV6	genome polyprotein
27	37	71.2	3391	1 GNVW26	genome polyprotein
28	37	71.2	3391	1 GNVWJA	genome polyprotein
29	37	71.2	3391	2 JS0219	polyprotein - deng

30	37	71.2	3432	1 GNVWJS	genome polyprotein
31	37	71.2	3432	1 GNVWJE	genome polyprotein
32	37	71.2	3434	1 GNVWVW	genome polyprotein
33	36	69.2	68	2 S20921	1-aminocyclopropan
34	36	69.2	532	2 B82102	membrane-bound lye
35	36	69.2	555	2 JQ1405	genome polyprotein
36	36	69.2	775	2 A47311	polyprotein(C, E,
37	36	69.2	792	2 C32401	genome polyprotein
38	36	69.2	792	2 B32401	genome polyprotein
39	36	69.2	792	2 A32401	genome polyprotein
40	36	69.2	1226	1 GNVWVP	genome polyprotein
41	36	69.2	1440	1 GNVWJP	genome polyprotein
42	36	69.2	3390	1 GNVWV3	genome polyprotein
43	36	69.2	3396	1 A42551	genome polyprotein
44	35	67.3	129	2 A42692	T-cell receptor al
45	35	67.3	130	2 A31211	T-cell receptor al
46	35	67.3	208	2 T33341	hypothetical prote
47	35	67.3	287	1 RWMSC8	T-cell receptor al
48	35	67.3	286	2 A91131	tagatose-1,6-bisph
49	35	67.3	286	2 A85976	tagatose-bisphosph
50	35	67.3	286	2 E65103	tagatose-bisphosph
51	35	67.3	286	2 T49369	hypothetical prote
52	35	67.3	442	2 S61165	hypothetical prote
53	35	67.3	449	2 G72393	hypothetical prote
54	35	67.3	455	2 S56695	1-aminocyclopropan
55	35	67.3	467	2 T10854	1-aminocyclopropan
56	35	67.3	472	1 T10889	1-aminocyclopropan
57	35	67.3	473	2 T16999	1-aminocyclopropan
58	35	67.3	476	1 S19679	1-aminocyclopropan
59	35	67.3	484	2 JC5779	4-carboxy-2-hydrox
60	35	67.3	485	1 S19677	1-aminocyclopropan
61	35	67.3	485	2 JW0056	1-aminocyclopropan
62	35	67.3	486	2 S10772	2-hydroxymuconic s
63	35	67.3	486	2 E42902	2-hydroxymuconic s
64	35	67.3	490	2 S31450	1-aminocyclopropan
65	35	67.3	491	2 T03978	1-aminocyclopropan
66	35	67.3	505	2 T31272	4-carboxy-2-hydrox
67	35	67.3	639	2 A32935	protein P1 - Entam
68	35	67.3	853	2 A71339	probable outer mem
69	35	67.3	1058	2 T08935	COP1-interacting p
70	35	67.3	1107	2 T20578	hypothetical prote
71	35	67.3	1119	2 T20577	hypothetical prote
72	35	67.3	1127	1 GNVWD2	genome polyprotein
73	35	67.3	1584	2 T00026	brain-specific ang
74	34	65.4	97	1 HMIWH6	hemagglutinin prec
75	34	65.4	166	2 S09223	membrane protein -
76	34	65.4	166	2 S09225	membrane protein -
77	34	65.4	191	2 F90392	hypothetical prote
78	34	65.4	210	1 JFBYA2	mating-type regula
79	34	65.4	222	2 AE2003	hypothetical prote
80	34	65.4	226	1 PWXL6	H+-transporting tw
81	34	65.4	287	2 E96756	hypothetical prote
82	34	65.4	312	2 T35111	probable tRNA delc
83	34	65.4	333	2 A75574	conserved esterase
84	34	65.4	365	2 B69114	1-aminocyclopropan
85	34	65.4	465	2 S54012	1-aminocyclopropan
86	34	65.4	465	2 S56176	1-aminocyclopropan
87	34	65.4	469	1 A57540	1-aminocyclopropan
88	34	65.4	470	2 T46036	1-aminocyclopropan
89	34	65.4	475	2 JQ2214	1-aminocyclopropan
90	34	65.4	566	1 HMVISA	hemagglutinin prec
91	34	65.4	597	2 F90481	conserved hypother
92	34	65.4	623	2 T48859	disease resistance
93	34	65.4	623	2 T06674	hypothetical prote
94	34	65.4	629	2 G84481	probable receptor-
95	34	65.4	665	2 PS0043	genome polyprotein
96	34	65.4	961	2 T03467	NADH dehydrogenase
97	34	65.4	1041	2 C87645	AcB/AcrD/AcrF fam
98	34	65.4	1057	2 T16676	hypothetical prote
99	33	63.5	68	2 S20920	1-aminocyclopropan
100	33	63.5	135	2 PC4252	hypothetical 135 p

ALIGNMENTS

RESULT 1  
A:7666  
structural polyprotein - dengue virus type 4 (fragment)  
N:Contains: capsid protein; envelope glycoprotein; membrane protein precursor  
C:Species: dengue virus type 4  
C>Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 31-Dec-2004  
C:Accession: A47666  
R;Kawano, H.; Rostapshov, V.; Rosen, L.; Lai, C.J.  
J. Virol. 67, 6567-6575, 1993  
A:Title: Genetic determinants of dengue type 4 virus neurovirulence for mice.  
A:Reference number: A47666; MUID:94016840; PMID:8411360  
A:Accession: A47666  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-773 <KAW>  
A:Cross-references: UNIPROT:O86654; UNIPARC:UPI0000F8175; GB:S66064; NID:G432575; PIDN:  
A:Experimental source: H241-P  
A:Note: sequence extracted from NCBI backbone (NCBI:138430, NCBI:138431)  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: glycoprotein; polyprotein

Query Match 88.5%; Score 46; DB 2; Length 773;  
Best Local Similarity 88.9%; Pred. No. 1.9; Mismatches 0; Gaps 0;  
Matches 8; Conservative 0; Indels 1; Indels 0; Gaps 0;  
QY 1 VESWFLRNP 9  
|||||  
DB 236 VESWILRNP 244

RESULT 2  
GNWVDF  
genome polyprotein - dengue virus type 4  
N:Contains: capsid protein; envelope protein; membrane protein; nonstructural protein 5;  
nonstructural protein NS4a; nonstructural protein NS4b  
C:Species: dengue virus type 4  
C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 31-Dec-2004  
C:Accession: A94352; A94364; A26897; A29121  
R;Zhao, B.; Mackow, E.; Buckler-White, A.; Markoff, L.; Chanock, R.M.; Lai, C.J.; Makino  
Virolgy 155, 77-88, 1986  
A:Title: Cloning full-length dengue type 4 viral DNA sequences: analysis of genes coding  
A:Reference number: A94352; MUID:87044106; PMID:3022479  
A:Accession: A94352  
A:Molecule type: Genomic RNA  
A:Residues: 1-776 <ZHA>  
A:Cross-references: UNIPARC:UPI0000174A08; GB:M14931  
R;Mackow, E.; Makino, Y.; Zhao, B.; Zhang, Y.M.; Markoff, L.; Buckler-White, A.; Guiler,  
Virolgy 159, 217-228, 1987  
A:Title: The nucleotide sequence of dengue type 4 virus: analysis of genes coding for non  
A:Reference number: A94364; MUID:87293881; PMID:3039728  
A:Accession: A94364  
A:Molecule type: Genomic RNA  
A:Residues: 774-3386 <MAC>  
A:Cross-references: UNIPARC:UPI0000174A09; GB:M17255  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F:2-113/Product: capsid protein #status predicted <CAP>  
F:42-58/Domain: transmembrane #status predicted <TM1>  
F:100-116/Domain: transmembrane #status predicted <TM2>  
F:114-279/Product: membrane protein precursor #status predicted <MEP>  
F:114-204/Domain: nonterminal signal sequence #status predicted <SIG>  
F:205-279/Product: membrane protein #status predicted <MBM>  
F:267-283/Domain: transmembrane #status predicted <TM3>  
F:280-773/Product: envelope protein #status predicted <ENV>  
F:728-744/Domain: transmembrane #status predicted <TM4>  
F:753-769/Domain: transmembrane #status predicted <TM5>  
F:774-1184/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1157-1179/Domain: transmembrane #status predicted <TM6>  
F:1185-1343/Product: nonstructural protein NS2a #status predicted <N2A>  
F:1344-1473/Product: nonstructural protein NS2b #status predicted <N2B>

F:1474-2091/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1666-1673/Region: nucleotide-binding motif A (P-loop)  
F:1753-1758/Region: nucleotide-binding motif B  
F:1757-1760/Region: DEAH motif  
F:2092-2374/Product: nonstructural protein NS4a #status predicted <N4A>  
F:2375-2486/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2487-3386/Product: nonstructural protein NS5 #status predicted <NS5>  
F:1182,346,432,750,903,980,2296,2300,2341,2382,2452,2582,2639,2699,2719,2913,3310/Binding  
Query Match 88.5%; Score 46; DB 1; Length 3386;  
Best Local Similarity 88.9%; Pred. No. 8.9; Mismatches 0; Gaps 0;  
Matches 8; Conservative 0; Indels 1; Indels 0; Gaps 0;  
QY 1 VESWFLRNP 9  
|||||  
DB 236 VESWILRNP 244

RESULT 3  
GNWVS  
genome polyprotein - St. Louis encephalitis virus (strain MS1-7) (fragment)  
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural  
C:Species: St. Louis encephalitis virus  
C>Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 31-Dec-2004  
C:Accession: A27531  
R;Trent, D.W.; Kinney, R.M.; Johnson, B.J.B.; Vorndam, A.V.; Grant, J.A.; Deubel, V.; Ric  
Virolgy 156, 293-304, 1987  
A:Title: Partial nucleotide sequence of St. Louis encephalitis virus RNA: structural prot  
A:Reference number: A27531; MUID:87122172; PMID:3027980  
A:Accession: A27531  
A:Molecule type: genomic RNA  
A:Residues: 1-1525 <TRE>  
A:Cross-references: UNIPARC:UPI0000174A04; GB:M16614; NID:G334865  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; poly  
F:1-119/Product: capsid protein C #status predicted <CPC>  
F:108-119/Domain: transmembrane #status predicted <TM1>  
F:120-288/Product: envelope protein M #status predicted <EPP>  
F:214-288/Product: envelope protein M #status predicted <EPM>  
F:253-268/Domain: transmembrane #status predicted <TM2>  
F:274-288/Domain: transmembrane #status predicted <TM3>  
F:289-789/Product: major envelope protein E #status predicted <EPE>  
F:751-763/Domain: transmembrane #status predicted <TM4>  
F:768-787/Domain: transmembrane #status predicted <TM5>  
F:790-1203/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1173-1188/Domain: transmembrane #status predicted <TM6>  
F:1204-1368/Product: nonstructural protein NS2a #status predicted <NSA>  
F:1369-1499/Product: nonstructural protein NS2b #status predicted <NSB>  
F:1500-1525/Product: nonstructural protein NS3 #status predicted <NS3>  
F:136,269,442,602,919,964,996,1189/Binding site: carbonylate (Asn) (covalent) #status p

Query Match 84.6%; Score 44; DB 1; Length 1525;  
Best Local Similarity 77.8%; Pred. No. 8.9; Mismatches 1; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Indels 1; Indels 0; Gaps 0;  
QY 1 VESWFLRNP 9  
|||||  
DB 245 VESWFLRNP 253

RESULT 4  
GNWVY  
genome polyprotein - yellow fever virus (strain 17D)  
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: yellow fever virus  
C>Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 31-Dec-2004  
C:Accession: A03914  
R;Rice, C.M.; Lencches, E.M.; Eddy, S.R.; Shin, S.J.; Sheets, R.L.; Strauss, J.H.  
Science 229, 726-733, 1985  
A:Title: Nucleotide sequence of yellow fever virus: implications for flavivirus gene exp  
A:Reference number: A03914; MUID:85272570; PMID:4023707  
A:Accession: A03914



A:Molecule type: genomic RNA  
A;Residues: 1-3411 <RIC>  
A;Cross-references: UNIPROT:P03314; UNIPARC:UPI0000131E82; GB:X03700; GB:K02749; NID:G59  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F:2-210/Product: capsid protein C #status predicted <CPC>  
F:211-285/Product: envelope protein M #status predicted <EPM>  
F:249-289/Product: transmembrane #status predicted <TM1>  
F:271-285/Domain: transmembrane #status predicted <TM2>  
F:286-778/Product: major envelope protein E #status predicted <MEE>  
F:740-753/Domain: transmembrane #status predicted <TM3>  
F:755-778/Domain: transmembrane #status predicted <TM4>  
F:775-1187/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1159-1180/Domain: transmembrane #status predicted <TM5>  
F:1188-1354/Product: nonstructural protein NS2a #status predicted <N2A>  
F:1355-1484/Product: nonstructural protein NS2b #status predicted <N2B>  
F:1485-2107/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1682-1689/Region: nucleotide-binding motif A (P-loop)  
F:1769-1774/Region: nucleotide-binding motif A (P-loop)  
F:1773-1776/Region: DEAH motif  
F:2108-2394/Product: nonstructural protein NS4a #status predicted <N4A>  
F:2395-2506/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2507-3411/Product: nonstructural protein NS5 #status predicted <NS5>  
F:134, 150, 172, 266, 594, 755, 908, 986, 1796, 2062, 2320, 2346, 2408, 2467, 2720, 2734, 2740/Binding site: car  
Query Match 82.7%; Score 43; DB 1; Length 3411;  
Best Local Similarity 66.7%; Pred. No. 32;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 VESWFLRNP 9  
Db 242 IERWVRNP 250  
RESULT 5  
GNWYVP  
Genome polyprotein - yellow fever virus (strain Pasteur 17D-204)  
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: Yellow fever virus  
C;Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 31-Dec-2004  
C;Accession: S07757  
R;Dupuy, A.; Despres, P.; Cahour, A.; Girard, M.; Bouloy, M.  
Nucleic Acids Res. 17, 3989, 1989  
A;Title: Nucleotide sequence comparison of the genome of two 17D-204 yellow fever vaccin  
A;Reference number: S07757; MUID:89282413; PMID:2734112  
A;Accession: S07757  
A;Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: genomic RNA  
A;Residues: 1-3411 <DUP>  
A;Cross-references: UNIPROT:P19901; UNIPARC:UPI0000131E83; EMBL:X15062; NID:G62289; PIDN  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1989, in c  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;  
F:2-210/Product: capsid protein C #status predicted <CPC>  
F:105-125/Domain: transmembrane #status predicted <TM1>  
F:211-285/Product: envelope protein M #status predicted <EPM>  
F:271-289/Domain: transmembrane #status predicted <TM2>  
F:286-778/Product: major envelope protein E #status predicted <MEE>  
F:736-753/Domain: transmembrane #status predicted <TM3>  
F:756-778/Domain: transmembrane #status predicted <TM4>  
F:779-1187/Product: nonstructural protein NS1 #status predicted <NS1>  
F:1133-1151/Domain: transmembrane #status predicted <TM5>  
F:1160-1175/Domain: transmembrane #status predicted <TM6>  
F:1188-1354/Product: nonstructural protein NS2a #status predicted <N2A>  
F:1355-1484/Product: nonstructural protein NS2b #status predicted <N2B>  
F:1485-2107/Product: nonstructural protein NS3 #status predicted <NS3>  
F:1682-1689/Region: nucleotide-binding motif A (P-loop)  
F:1769-1774/Region: nucleotide-binding motif A (P-loop)  
F:1773-1776/Region: DEAH motif  
F:2108-2394/Product: nonstructural protein NS4a #status predicted <N4A>  
F:2395-2506/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2507-3411/Product: nonstructural protein NS5 #status predicted <NS5>

F:134, 150, 172, 594, 908, 986, 1796, 2062, 2320, 2346, 2408, 2467, 2720, 2734, 2740/Binding site: car  
Query Match 82.7%; Score 43; DB 1; Length 3411;  
Best Local Similarity 66.7%; Pred. No. 32;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 VESWFLRNP 9  
Db 242 IERWVRNP 250  
RESULT 6  
GNWVWV  
Genome polyprotein - West Nile virus  
N;Contains: core protein V2; membrane-associated glycoprotein NV2 precursor; membrane-as  
sor; nonstructural protein NV5  
C;Species: West Nile virus  
C;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 31-Dec-2004  
C;Accession: A25256  
R;Castle, E.; Leidner, U.; Nowak, T.; Wengler, G.; Wengler, G.  
Virology 149, 10-26, 1986  
A;Title: Primary structure of the West Nile flavivirus genome region coding for all nons  
A;Reference number: A25256; MUID:86124703; PMID:3753811  
A;Accession: A25256  
A:Molecule type: genomic RNA  
A;Residues: 1-3430 <CAS>  
A;Cross-references: UNIPROT:P06935; UNIPARC:UPI0000131E81; GB:M10103; GB:M12294; NID:G33  
A;Note: parts of this sequence, including the amino ends of the mature proteins, were de  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; core protein; glycoprotein; membrane-associated protein; nucleotide bin  
F:1-92/Product: core protein V2 #status predicted <CV2>  
F:105-233/Product: membrane-associated glycoprotein NV2 precursor #status predicted <NV2>  
F:124-233/Domain: nonterminal signal sequence #status predicted <NS>  
F:124-233/Product: membrane-associated glycoprotein NV2 #status predicted <NV2>  
F:216-233/Product: membrane-associated nonglycosylated protein V1 #status predicted <NV1>  
F:275-290/Domain: membrane-associated glycoprotein V3 precursor #status predicted <NV3>  
F:291-787/Product: nonterminal signal sequence #status predicted <NS>  
F:788-2109/Product: membrane-associated glycoprotein V3 #status predicted <NV3>  
F:788-2109/Product: membrane-associated glycoprotein V4 #status predicted <NV4>  
F:1695-1702/Region: nucleotide-binding motif A (P-loop)  
F:1782-1787/Region: nucleotide-binding motif B  
F:1786-1789/Region: DEAH motif  
F:2580-3427/Product: nonstructural protein NV5 #status predicted <NV5>  
F:138, 917, 962, 994, 1289, 1659, 1966, 2336, 2459, 2489, 2573, 2739, 2759, 2864, 2902/Binding site: ca  
Query Match 80.8%; Score 42; DB 1; Length 3430;  
Best Local Similarity 87.5%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 2 ESWFLRNP 9  
Db 248 ESWILRNP 255  
RESULT 7  
GNWVWV  
Genome polyprotein - Kunjin virus (strain MRM61C)  
N;Contains: capsid protein C; envelope protein E; membrane protein M; nonstructural prote  
in NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: Kunjin virus  
C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 31-Dec-2004  
C;Accession: A28697  
R;Coia, G.; Parker, M.D.; Speight, G.; Byrne, M.E.; Westaway, E.G.  
J. Gen. Virol. 69, 1-21, 1988  
A;Title: Nucleotide and complete amino acid sequences of Kunjin virus: definitive gene o  
A;Reference number: A28697; MUID:88099524; PMID:2826659  
A;Accession: A28697  
A:Molecule type: genomic RNA  
A;Residues: 1-3433 <COI>  
A;Cross-references: UNIPROT:P14335; UNIPARC:UPI0000131E43; GB:D00246; NID:G221966; PIDN:f  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; capsid protein; envelope protein; membrane protein; nonstructural protei  
F:2-123/Product: capsid protein C #status predicted <CPC>

F:124-290/Product: membrane protein M precursor #status predicted <MPP>  
 F:124-215/Domain: nonterminal signal sequence #status predicted <SIG>  
 F:216-290/Product: membrane protein M #status predicted <MPM>  
 F:291-791/Product: envelope protein E #status predicted <EPE>  
 F:792-1143/Product: nonstructural protein NS1 #status predicted <NS1>  
 F:1144-1374/Product: nonstructural protein NS2a #status predicted <N2a>  
 F:1375-1505/Product: nonstructural protein NS2b #status predicted <N2b>  
 F:1506-2124/Product: nonstructural protein NS3 #status predicted <NS3>  
 F:1699-1706/Product: nucleotide-binding motif A (P-loop)  
 F:1786-1791/Region: nucleotide-binding motif B  
 F:1790-1793/Region: DEAH motif  
 F:2125-2273/Product: nonstructural protein NS4a #status predicted <N4a>  
 F:2274-2528/Product: nonstructural protein NS4b #status predicted <N4b>  
 F:2529-3433/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 80.8%; Score 42; DB 1; Length 3433;  
 Best Local Similarity 87.5%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ESWFLRNP 9  
 |||||  
 Db 248 ESWILRNP 255

RESULT 8  
 D69104  
 conserved hypothetical protein MTH1778 - Methanobacterium thermoautotrophicum (strain De  
 C;Species: Methanobacterium thermoautotrophicum  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C;Accession: D69104  
 R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
 i; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
 A;Reference number: A69000; MUID:98037514; PMID:9371463  
 A;Accession: D69104  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-73 <MTH>  
 A;Cross-references: UNIPROT:O27806; UNIPARC:UPI00000665CF; GB:AE000666; NID  
 A;Experimental source: strain Delta H  
 C;Genetics:  
 A;Gene: MTH1778  
 A;Start codon: GTG

Query Match 76.9%; Score 40; DB 2; Length 73;  
 Best Local Similarity 85.7%; Pred. No. 2;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLRNP 9  
 |||||  
 Db 33 SWFVRNP 39

RESULT 9  
 H69453  
 hypothetical protein AF1633 - Archaeoglobus fulgidus  
 C;Species: Archaeoglobus fulgidus  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C;Accession: H69453  
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.P.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.P.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A;Reference number: A69250; MUID:98049343; PMID:9389475  
 A;Accession: H69453  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-227 <KLE>

A;Cross-references: UNIPROT:O28640; UNIPARC:UPI0000056CB; GB:AE000989; GB:AE000782; NID:  
 Query Match 75.0%; Score 39; DB 2; Length 227;  
 Best Local Similarity 85.7%; Pred. No. 9.8;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SWFLRNP 9  
 |||||  
 Db 161 SWFLQNP 167

RESULT 10  
 S58857  
 botulinum neurotoxin type B hemagglutinin component, 33K - Clostridium botulinum (strain  
 N/Alternate names: protein HA-33  
 C;Species: Clostridium botulinum  
 A;Variety: strain Eklund 17B  
 C;Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Jul-2004  
 C;Accession: S58857  
 R;East, A.K.; Stacey, J.M.; Collins, M.D.  
 Syst. Appl. Microbiol. 17, 306-312, 1994  
 A;Title: Cloning and sequencing of a hemagglutinin component of the botulinum neurotoxin  
 A;Reference number: S58855  
 A;Accession: S58857  
 A;Molecule type: DNA  
 A;Residues: 1-232 <EAS>  
 A;Cross-references: UNIPROT:Q45968; UNIPARC:UPI00000BBE2B; EMBL:X79103; NID:g870932; PID:  
 A;Experimental source: strain Eklund 17B  
 C;Keywords: hemagglutinin; neurotoxin

Query Match 75.0%; Score 39; DB 2; Length 292;  
 Best Local Similarity 62.5%; Pred. No. 13;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ESWFLRNP 9  
 :||:|  
 Db 285 QKWFI RNP 292

RESULT 11  
 F84807  
 hypothetical protein At2g38650 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 31-Dec-2004  
 C;Accession: F84807  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: F84807  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-582 <STO>  
 A;Cross-references: UNIPROT:Q9ZVI7; UNIPARC:UPI000017989F; GB:AE002093; NID:G3786007; PII  
 C;Genetics:  
 A;Gene: At2g38650  
 A;Map position: 2

Query Match 75.0%; Score 39; DB 2; Length 582;  
 Best Local Similarity 55.6%; Pred. No. 26;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9  
 :||:|  
 Db 331 MKQWFI RNP 339

RESULT 12  
 T29563  
 hypothetical protein T12E12.5 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T29563  
 R;Bradshaw, H.; Stellyes, L.  
 A;Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 A;Description: The sequence of C. elegans cosmid T12E12.  
 A;Reference number: Z20641  
 A;Accession: T29563  
 A;Status: preliminary; translated from GB/EMBL/DBDJ  
 A;Molecule type: DNA  
 A;Residues: 1-94 <BRA>  
 A;Cross-references: UNIPARC:UPI00001641F8; EMBL:U61944; PIDN:AB03122.1; GSPDB:GN000022;  
 A;Experimental source: strain Bristol N2; clone T12E12  
 C;Genetics:  
 A;Gene: CESP:T12E12.5  
 A;Map position: 4  
 A;Introns: 63/3

Query Match 71.2%; Score 37; DB 2; Length 94;  
 Best Local Similarity 85.7%; Pred. No. 8.9;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESWFLRN 8  
 Db 88 KSWFLRN 94  
 :|||||  
 :|||||

RESULT 13  
 S40144  
 premembrane protein - dengue virus type 2  
 C;Species: dengue virus type 2  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 31-Dec-2004  
 C;Accession: S40144  
 R;Shiu, S.Y.W.  
 A;Note: submitted to the EMBL Data Library, May 1993  
 A;Reference number: S40144  
 A;Accession: S40144  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-166 <SHI>  
 A;Cross-references: UNIPROT:Q66346; UNIPARC:UPI00000F6DD9; EMBL:X72849; NID:g437772; PID  
 C;Superfamily: hepatitis C virus genome polyprotein

Query Match 71.2%; Score 37; DB 2; Length 166;  
 Best Local Similarity 55.6%; Pred. No. 16;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
 Db 123 IETWILRHP 131  
 :|||:|  
 :|||:|

RESULT 14  
 S67182  
 hypothetical protein YOR280c - yeast (Saccharomyces cerevisiae)  
 N;Alternate names: hypothetical protein 05471  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 31-Dec-2004  
 C;Accession: S67182; S72050  
 R;Cheret, G.; Sor, F.  
 A;Note: submitted to the Protein Sequence Database, July 1996  
 A;Reference number: S67169  
 A;Accession: S67182  
 A;Molecule type: DNA  
 A;Residues: 1-266 <CHE>  
 A;Cross-references: UNIPROT:Q99369; UNIPARC:UPI000006B767; EMBL:275188; NID:gl420624; PI  
 A;Experimental source: strain S288C  
 R;Cheret, G.; Bernardi, A.; Sor, F.  
 A;Title: DNA sequence analysis of the VP1-SNF2 region on chromosome XV of Saccharomyces  
 A;Reference number: S72039; MUID:97051594; PMID:8896271  
 A;Accession: S72050  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA

A;Residues: 1-266 <CHW>  
 A;Cross-references: UNIPARC:UPI000006B767; EMBL:X89633; NID:gl279694; PIDN:CAA61785.1; P  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995  
 C;Genetics:  
 A;Cross-references: SGD:S0005806  
 A;Map position: 15R  
 A;Note: YOR280c  
 C;Superfamily: uncharacterized conserved protein

Query Match 71.2%; Score 37; DB 2; Length 266;  
 Best Local Similarity 66.7%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
 Db 76 VYGWFFRNP 84  
 :|||||  
 :|||||

RESULT 15  
 T06024  
 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14) T28119.50 - Arabidopsis thaliana  
 N;Alternate names: protein T28119.50  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
 C;Accession: T06024  
 R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba  
 A;Note: submitted to the Protein Sequence Database, March 1999  
 A;Reference number: Z15484  
 A;Accession: T06024  
 A;Molecule type: DNA  
 A;Residues: 1-469 <BEV>  
 A;Cross-references: UNIPROT:Q9T065; UNIPARC:UPI00000A09EF; EMBL:AL035709; GSPDB:GN000062;  
 A;Experimental source: cultivar Columbia; BAC clone T28119  
 C;Genetics:  
 A;Gene: ATSP:T28119.50  
 A;Map position: 4  
 A;Introns: 49/3; 93/3; 147/2  
 C;Superfamily: 1-aminocyclopropane-1-carboxylate synthase  
 C;Keywords: carbon-sulfur lyase; ethylene biosynthesis; phosphoprotein; pyridoxal phosph  
 P;272/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 71.2%; Score 37; DB 2; Length 469;  
 Best Local Similarity 55.6%; Pred. No. 48;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
 Db 55 IESWLAKNP 63  
 :||||:|  
 :||||:|

RESULT 16  
 S71174  
 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14) ACS5 - Arabidopsis thaliana  
 N;Alternate names: protein F6H11.90  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004  
 C;Accession: S71174; T05890; H46376  
 R;Liang, X.; Shen, N.F.; Theologis, A.  
 A;Note: submitted to the EMBL Data Library, February 1996  
 A;Description: Li+ regulated 1-aminocyclopropane-1-carboxylate synthase gene expression  
 tein kinases.  
 A;Reference number: S71174  
 A;Accession: S71174  
 A;Molecule type: mRNA  
 A;Residues: 1-470 <LIA>  
 A;Cross-references: UNIPROT:Q37001; UNIPROT:O49537; UNIPARC:UPI000009968D; EMBL:L29261; N  
 R;Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Schafke, M.; Schon, O.; Bancroft, I.; Mewes  
 A;Note: submitted to the Protein Sequence Database, April 1998  
 A;Reference number: Z15456  
 A;Accession: T05890  
 A;Molecule type: DNA  
 A;Residues: 1-470 <BEV>  
 A;Cross-references: UNIPARC:UPI000009968D; EMBL:AL021684; GSPDB:GN000063; ATSP:F6H11.90

A:Experimental source: cultivar Columbia; BAC clone F6H11  
R:Liang, X.; Abel, S.; Keller, J.A.; Shen, N.F.; Theologos, A.  
Proc. Natl. Acad. Sci. U.S.A. 89, 11046-11050, 1992  
A:Title: The 1-aminocyclopropane-1-carboxylate synthase gene family of Arabidopsis thaliana  
A:Reference number: A46376; MUID:93066381; PMID:1438312  
A:Accession: H46376  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 50-81 <BI>  
A:Cross-references: UNIPARC:UPI00000A1204  
A>Note: sequence extracted from NCBI backbone (NCBIP:118952)  
C:Genetics:  
A:Gene: ATSP:F6H11.90; ACS5; ACC5  
A:Map position: 5  
A:Introns: 93/3; 147/2  
C:Superfamily: 1-aminocyclopropane-1-carboxylate synthase  
C:Keywords: carbon-sulfur lyase; ethylene biosynthesis; phosphoprotein; pyridoxal phosphate  
F:727/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 71.2%; Score 37; DB 2; Length 470;  
Best Local Similarity 55.8%; Pred. No. 49;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9  
Db 55 IESWLTKNP 63

RESULT 17  
H91109  
hypothetical protein EC63848 [imported] - Escherichia coli (strain O157:H7, substrain R157:H7)  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: H91109  
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: H91109  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-482 <HAY>  
A:Cross-references: UNIPROT:Q8XBZ1; UNIPARC:UPI00000D0537; GB:BA000007; PIDN:BA037271.1;  
A:Experimental source: strain O157:H7, substrain R157:H7  
C:Genetics:  
A:Gene: EC63848

Query Match 71.2%; Score 37; DB 2; Length 482;  
Best Local Similarity 85.7%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ESWFLRN 8  
Db 471 ESWFFRN 477

RESULT 18  
B85955  
unknown protein encoded by ISEC8 [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: B85955  
R:Ferns, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: B85955  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-482 <STO>  
A:Cross-references: UNIPROT:Q8XBZ1; UNIPARC:UPI00000D0537; GB:AB005174; NID:g12517523; H

A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: 24317

Query Match 71.2%; Score 37; DB 2; Length 482;  
Best Local Similarity 85.7%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ESWFLRN 8  
Db 471 ESWFFRN 477

RESULT 19  
S25002  
1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14) - soybean  
C:Species: Glycine max (soybean)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S25002  
R:Li, D.; Li, N.; Mattoo, A.K.  
submitted to the EMBL Data Library, June 1992  
A:Description: Nucleotide sequence of soybean ACC synthase.  
A:Reference number: S25002  
A:Accession: S25002  
A:Molecule type: mRNA  
A:Residues: 1-484 <LIU>  
A:Cross-references: UNIPROT:P31531; UNIPARC:UPI0000124E35; EMBL:X67100; NID:g18557; PIDN:  
C:Superfamily: 1-aminocyclopropane-1-carboxylate synthase  
C:Keywords: carbon-sulfur lyase; ethylene biosynthesis; phosphoprotein; pyridoxal phosphate  
F:279/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 71.2%; Score 37; DB 2; Length 484;  
Best Local Similarity 66.7%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9  
Db 63 VEDWILNRP 71

RESULT 20  
S26214  
1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14) (clone pAIM-1) - mung bean  
N:Alternate names: ACC synthase  
C:Species: Vigna radiata (mung bean)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S26214; S26213; S20919  
R:Botella, J.R.; Arteca, J.M.; Schlagnhauser, C.D.; Arteca, R.N.; Phillips, A.T.  
Plant Mol. Biol. 20, 425-436, 1992  
A:Title: Identification and characterization of a full-length cDNA encoding for an auxin-  
of its mRNA in response to indole-3-acetic acid.  
A:Reference number: S26213; MUID:93043033; PMID:1421146  
A:Accession: S26214  
A:Molecule type: mRNA  
A:Residues: 1-484 <BOT>  
A:Cross-references: UNIPROT:Q43858; UNIPARC:UPI00000A0DAC; EMBL:Z11613; NID:g22069; PIDN:  
A:Experimental source: clone pAIM-1  
A:Accession: S26213  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 50-415, 'P', 417 <BOT>  
A:Cross-references: UNIPARC:UPI0000124E34; EMBL:Z11562; NID:g22067; PIDN:CAA77655.1; PID  
A:Experimental source: clone pAIM-1  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1992  
R:Botella, J.R.; Schlagnhauser, C.D.; Arteca, R.N.; Phillips, A.T.  
Plant Mol. Biol. 18, 793-797, 1992  
A:Title: Identification and characterization of three putative genes for 1-aminocyclopropan  
A:Reference number: S20919; MUID:92216056; PMID:1558953  
A:Accession: S20919  
A:Molecule type: DNA  
A:Residues: 27-94 <BOF>  
A:Cross-references: UNIPARC:UPI00000A4C08; GB:M80554; NID:g170628; PIDN:AAA53297.1; PID:  
A:Experimental source: Rwilcz cv. Berken, etiolated hypocotyls; clone pMAC-1

C;Superfamily: 1-aminocyclopropane-1-carboxylate synthase  
C;Keywords: carbon-sulfur lyase; ethylene biosynthesis; phosphoprotein; pyridoxal phosph  
F;279/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 71.2%; Score 37; DB 2; Length 484;  
Best Local Similarity 66.7%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
|||  
Db 63 VEDWILNRP 71

RESULT 21  
JQ1404  
genome polyprotein - dengue virus type 2 (strain TH-36) (fragment)  
N;Contains: envelope protein E; membrane-associated protein M; nonstructural protein NS1  
C;Species: dengue virus type 2  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-2004  
C;Accession: JQ1404  
R;Shiu, S.Y.W.; Jiang, W.R.; Porterfield, J.S.; Gould, E.A.  
J. Gen. Virol. 73, 207-212, 1992  
A;Title: Envelope protein sequences of dengue virus isolates TH-36 and TH-Sman, and iden  
A;Reference number: JQ1404; MUID:92113574; PMID:1339466  
A;Accession: JQ1404  
A;Molecule type: genomic RNA  
A;Residues: 1-555 <SHI>  
A;Cross-references: UNIPROT:P29984; UNIPARC:UPI0000131DF8; GB:D10514; DDBJ:D01074; NID:9  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: envelope protein; glycoprotein; nonstructural protein; polyprotein; transmem  
F;1-49/Product: transmembrane-associated protein M (fragment) #status predicted <MEM>  
F;37-53/Domain: transmembrane #status predicted <TM1>  
F;50-544/Product: envelope protein E #status predicted <ENV>  
F;496-512/Domain: transmembrane #status predicted <TM2>  
F;526-542/Domain: transmembrane #status predicted <TM3>  
F;545-555/Product: nonstructural protein NS1 (fragment) #status predicted <NON>  
F;116,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.2%; Score 37; DB 2; Length 555;  
Best Local Similarity 55.6%; Pred. No. 58;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
:|||  
Db 6 IETWILRHP 14

RESULT 22  
A48644  
polyprotein - dengue virus type 2 (strain Mexican) (fragment)  
C;Species: dengue virus type 2  
C;Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 31-Dec-2004  
C;Accession: A48644  
R;Ruiz, B.H.; Sanchez, I.; Ortega, G.J.; Lopez, I.; Ortiz-Ortiz, L.  
submitted to GenBank, October 1992  
A;Description: Nucleotide sequence and deduced amino-acid sequence of the structural pro  
A;Reference number: A48644  
A;Accession: A48644  
A;Status: preliminary  
A;Molecule type: genomic RNA  
A;Residues: 1-775 <RUI>  
A;Cross-references: UNIPROT:Q66398; UNIPARC:UPI000000EEB45; GB:L04561; NID:G323652; PIDN:  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: polyprotein

Query Match 71.2%; Score 37; DB 2; Length 775;  
Best Local Similarity 55.6%; Pred. No. 82;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
:|||  
Db 237 IETWILRHP 245

RESULT 23  
I50804  
polyprotein - Japanese encephalitis virus (fragment)  
C;Species: Japanese encephalitis virus  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 31-Dec-2004  
C;Accession: I50804  
R;Ni, H.; Barrett, A.D.  
J. Gen. Virol. 76, 401-407, 1995  
A;Title: Nucleotide and deduced amino acid sequence of the structural protein genes of J  
A;Reference number: I50804; MUID:95146981; PMID:7844559  
A;Accession: I50804  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-789 <NIX>  
A;Cross-references: UNIPROT:Q82865; UNIPARC:UPI000000F3E2D; EMBL:U03693; NID:G517403; PID:  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: polyprotein

Query Match 71.2%; Score 37; DB 2; Length 789;  
Best Local Similarity 62.5%; Pred. No. 84;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ESWFLRNP 9  
:|||  
Db 252 ENWIIRNP 259

RESULT 24  
GNWY8  
genome polyprotein - yellow fever virus (strain 1899/81) (fragment)  
N;Contains: amino end of nonstructural protein NS1; capsid protein C; envelope protein M  
C;Species: yellow fever virus  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 31-Dec-2004  
C;Accession: JU0374  
R;Ballinger-Crabtree, M.E.; Miller, B.R.  
J. Gen. Virol. 71, 2115-2121, 1990  
A;Title: Partial nucleotide sequence of South American yellow fever virus strain 1899/81  
A;Reference number: JU0374; MUID:91011358; PMID:2145394  
A;Accession: JU0374  
A;Molecule type: genomic RNA  
A;Residues: 1-1163 <BAL>  
A;Cross-references: UNIPROT:P29165; UNIPARC:UPI0000131E84; GB:D14458; GB:D00739; NID:G22  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; nucle  
F;2-121/Product: capsid protein C #status predicted <CAP>  
F;106-122/Domain: transmembrane #status predicted <TM1>  
F;122-285/Product: envelope protein M #status predicted <PRM>  
F;251-267/Domain: transmembrane #status predicted <TM2>  
F;271-287/Domain: transmembrane #status predicted <TM3>  
F;286-778/Product: major envelope protein E #status predicted <ENP>  
F;733-753/Domain: transmembrane #status predicted <TM4>  
F;732-778/Domain: transmembrane #status predicted <TM5>  
F;779-1163/Product: nonstructural protein NS1 (fragment) #status predicted <NS1>  
F;1068-1075/Region: nucleotide-binding motif A (P-loop)  
F;1133-1151/Domain: transmembrane #status predicted <TM6>  
F;134,150,172,266,554,594,755,908,986/Binding site: carbohydrate (Asn) (covalent) #statu

Query Match 71.2%; Score 37; DB 1; Length 1163;  
Best Local Similarity 55.6%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9  
:|||  
Db 242 IERWLVRNP 250

RESULT 25  
GNWYDP  
genome polyprotein - dengue virus type 2 (strain PR159/S1)  
N;Contains: capsid protein; envelope protein; membrane protein; nonstructural protein NS  
a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: dengue virus type 2







F:1505-2123/Product: nonstructural protein NS3 #status predicted <NS3>  
 F:1698-1705/Region: nucleotide-binding motif A (P-loop)  
 F:1785-1790/Region: nucleotide-binding motif B  
 F:1789-1792/Region: DEAH motif  
 F:2124-2412/Product: nonstructural protein NS4a #status predicted <N4A>  
 F:2413-2527/Product: nonstructural protein NS4b #status predicted <N4B>  
 F:2528-3432/Product: nonstructural protein NS5 #status predicted <NS5>  
 F:142,448,924,1001,1594,1950,2463,2491,2761,2866,2904/Binding site: carbohydrate (Asn)

Query Match 71.2%; Score 37; DB 1; Length 3432;  
 Best Local Similarity 62.5%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ESWFLRNP 9  
 Db 252 ENWLRNP 259

Search completed: August 31, 2006, 11:51:52  
 Job time : 19.25 secs



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2006, 11:33:43 ; Search time 139 Seconds  
(without alignments)  
59.893 Million cell updates/sec

Title: DENGUE\_SEROTYPE4  
Perfect score: 52  
Sequence: 1 veswflnp 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 92501592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_7.2.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	88.5	131	2	Q8V1J4 dengue viru
2	46	88.5	131	2	Q8V1J5 dengue viru
3	46	88.5	131	2	Q8V1J6 dengue viru
4	46	88.5	131	2	Q8V1J7 dengue viru
5	46	88.5	131	2	Q8V1J9 dengue viru
6	46	88.5	131	2	Q8V1K0 dengue viru
7	46	88.5	131	2	Q8V1K1 dengue viru
8	46	88.5	348	2	Q3ZPL2 dengue viru
9	46	88.5	348	2	Q3ZPL4 dengue viru
10	46	88.5	348	2	Q3ZPL5 dengue viru
11	46	88.5	545	2	Q66432 dengue viru
12	46	88.5	545	2	Q66433 dengue viru
13	46	88.5	545	2	Q66435 dengue viru
14	46	88.5	545	2	Q66436 dengue viru
15	46	88.5	545	2	Q66437 dengue viru
16	46	88.5	545	2	Q66438 dengue viru
17	46	88.5	545	2	Q66439 dengue viru
18	46	88.5	545	2	Q66440 dengue viru
19	46	88.5	545	2	Q66441 dengue viru
20	46	88.5	545	2	Q66442 dengue viru
21	46	88.5	545	2	Q66443 dengue viru
22	46	88.5	545	2	Q66444 dengue viru
23	46	88.5	545	2	Q66445 dengue viru
24	46	88.5	545	2	Q66446 dengue viru
25	46	88.5	545	2	Q66447 dengue viru
26	46	88.5	545	2	Q66448 dengue viru
27	46	88.5	545	2	Q66449 dengue viru
28	46	88.5	583	2	Q3ZPL3 dengue viru
29	46	88.5	590	2	Q68SA7 dengue viru
30	46	88.5	646	2	Q2YED3 dengue viru
31	46	88.5	646	2	Q6YFK8 dengue viru

32	46	88.5	646	2	Q6YFK9_DEN4	Q6YFK9 dengue viru
33	46	88.5	646	2	Q6YFLO_DEN4	Q6YFLO dengue viru
34	46	88.5	646	2	Q80K23_DEN4	Q80K23 dengue viru
35	46	88.5	646	2	Q80K24_DEN4	Q80K24 dengue viru
36	46	88.5	646	2	Q80K25_DEN4	Q80K25 dengue viru
37	46	88.5	646	2	Q80K26_DEN4	Q80K26 dengue viru
38	46	88.5	646	2	Q80K27_DEN4	Q80K27 dengue viru
39	46	88.5	646	2	Q80K28_DEN4	Q80K28 dengue viru
40	46	88.5	646	2	Q80K29_DEN4	Q80K29 dengue viru
41	46	88.5	646	2	Q80L00_DEN4	Q80L00 dengue viru
42	46	88.5	646	2	Q80L01_DEN4	Q80L01 dengue viru
43	46	88.5	646	2	Q80L02_DEN4	Q80L02 dengue viru
44	46	88.5	646	2	Q80L03_DEN4	Q80L03 dengue viru
45	46	88.5	646	2	Q80L04_DEN4	Q80L04 dengue viru
46	46	88.5	646	2	Q80L05_DEN4	Q80L05 dengue viru
47	46	88.5	646	2	Q80L06_DEN4	Q80L06 dengue viru
48	46	88.5	646	2	Q80L07_DEN4	Q80L07 dengue viru
49	46	88.5	646	2	Q80L08_DEN4	Q80L08 dengue viru
50	46	88.5	646	2	Q80L09_DEN4	Q80L09 dengue viru
51	46	88.5	646	2	Q80L10_DEN4	Q80L10 dengue viru
52	46	88.5	646	2	Q80L11_DEN4	Q80L11 dengue viru
53	46	88.5	646	2	Q80L12_DEN4	Q80L12 dengue viru
54	46	88.5	646	2	Q80L13_DEN4	Q80L13 dengue viru
55	46	88.5	646	2	Q80L14_DEN4	Q80L14 dengue viru
56	46	88.5	646	2	Q80L15_DEN4	Q80L15 dengue viru
57	46	88.5	646	2	Q80L16_DEN4	Q80L16 dengue viru
58	46	88.5	646	2	Q80L17_DEN4	Q80L17 dengue viru
59	46	88.5	646	2	Q80L18_DEN4	Q80L18 dengue viru
60	46	88.5	678	2	Q6KEY4_DEN4	Q6KEY4 dengue viru
61	46	88.5	773	2	Q86654_DEN4	Q86654 dengue viru
62	46	88.5	774	2	Q5DPX7_DEN4	Q5DPX7 dengue viru
63	46	88.5	850	2	Q6YFQ8_DEN4	Q6YFQ8 dengue viru
64	46	88.5	850	2	Q6YFR2_DEN4	Q6YFR2 dengue viru
65	46	88.5	850	2	Q6YFR6_DEN4	Q6YFR6 dengue viru
66	46	88.5	850	2	Q6YFS0_DEN4	Q6YFS0 dengue viru
67	46	88.5	850	2	Q6YFS4_DEN4	Q6YFS4 dengue viru
68	46	88.5	850	2	Q6YFS8_DEN4	Q6YFS8 dengue viru
69	46	88.5	850	2	Q6YFT2_DEN4	Q6YFT2 dengue viru
70	46	88.5	850	2	Q6YFT6_DEN4	Q6YFT6 dengue viru
71	46	88.5	850	2	Q6YFU0_DEN4	Q6YFU0 dengue viru
72	46	88.5	850	2	Q6YFU4_DEN4	Q6YFU4 dengue viru
73	46	88.5	850	2	Q6YFU8_DEN4	Q6YFU8 dengue viru
74	46	88.5	850	2	Q6YFV2_DEN4	Q6YFV2 dengue viru
75	46	88.5	850	2	Q6YFV6_DEN4	Q6YFV6 dengue viru
76	46	88.5	850	2	Q6YFW0_DEN4	Q6YFW0 dengue viru
77	46	88.5	850	2	Q6YFW4_DEN4	Q6YFW4 dengue viru
78	46	88.5	850	2	Q6YFW8_DEN4	Q6YFW8 dengue viru
79	46	88.5	850	2	Q6YFX2_DEN4	Q6YFX2 dengue viru
80	46	88.5	850	2	Q6YFX6_DEN4	Q6YFX6 dengue viru
81	46	88.5	850	2	Q6YFY0_DEN4	Q6YFY0 dengue viru
82	46	88.5	850	2	Q6YFY4_DEN4	Q6YFY4 dengue viru
83	46	88.5	850	2	Q6YFY8_DEN4	Q6YFY8 dengue viru
84	46	88.5	850	2	Q6YFZ2_DEN4	Q6YFZ2 dengue viru
85	46	88.5	850	2	Q6YFZ6_DEN4	Q6YFZ6 dengue viru
86	46	88.5	850	2	Q6YG00_DEN4	Q6YG00 dengue viru
87	46	88.5	850	2	Q6YG04_DEN4	Q6YG04 dengue viru
88	46	88.5	850	2	Q6YG08_DEN4	Q6YG08 dengue viru
89	46	88.5	850	2	Q6YGL2_DEN4	Q6YGL2 dengue viru
90	46	88.5	850	2	Q6YGL6_DEN4	Q6YGL6 dengue viru
91	46	88.5	850	2	Q6YG20_DEN4	Q6YG20 dengue viru
92	46	88.5	850	2	Q6YG24_DEN4	Q6YG24 dengue viru
93	46	88.5	850	2	Q6YG28_DEN4	Q6YG28 dengue viru
94	46	88.5	850	2	Q6YG32_DEN4	Q6YG32 dengue viru
95	46	88.5	850	2	Q6YG36_DEN4	Q6YG36 dengue viru
96	46	88.5	850	2	Q6YG40_DEN4	Q6YG40 dengue viru
97	46	88.5	850	2	Q6YG44_DEN4	Q6YG44 dengue viru
98	46	88.5	850	2	Q6YG48_DEN4	Q6YG48 dengue viru
99	46	88.5	850	2	Q6YG52_DEN4	Q6YG52 dengue viru
100	46	88.5	850	2	Q6YG56_DEN4	Q6YG56 dengue viru

ALIGNMENTS

```

DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
FT NON_TER 1 131
FT NON_TER 131 131
SQ SEQUENCE 131 AA; 14975 MW; 10F1CD64F5364148 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 131;
Best Local Similarity 88.9%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 123 VESWILRNP 131

RESULT 3
Q8VLJ6_DEN4 PRELIMINARY; PRT; 131 AA.
AC Q8VLJ6;
DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2002, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Premembrane (Fragment).
GN Name=prM;
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22249823; PubMed=12363071;
RA Diaz F.J., Farfan-Ale J.A., Olson K.E., Llorono-Pino M.A., Gubler D.J.,
RA Blair C.D., Black W.C. IV, Beaty B.J.;
RT "Genetic variation within the premembrane coding region of dengue
RT viruses from the Yucatan peninsula of Mexico.";
RL Am. J. Trop. Med. Hyg. 67:93-101(2002).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AF459626; AAL67828.1; -; Genomic_RNA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
FT NON_TER 1 131
FT NON_TER 131 131
SQ SEQUENCE 131 AA; 14971 MW; 10F1CA7F98878148 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 131;
Best Local Similarity 88.9%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 123 VESWILRNP 131

RESULT 2
Q8VLJ5_DEN4 PRELIMINARY; PRT; 131 AA.
AC Q8VLJ5;
DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2002, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Premembrane (Fragment).
GN Name=prM;
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22249823; PubMed=12363071;
RA Diaz F.J., Farfan-Ale J.A., Olson K.E., Llorono-Pino M.A., Gubler D.J.,
RA Blair C.D., Black W.C. IV, Beaty B.J.;
RT "Genetic variation within the premembrane coding region of dengue
RT viruses from the Yucatan peninsula of Mexico.";
RL Am. J. Trop. Med. Hyg. 67:93-101(2002).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AF459626; AAL67828.1; -; Genomic_RNA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
FT NON_TER 1 131
FT NON_TER 131 131
SQ SEQUENCE 131 AA; 14971 MW; 10F1CA7F98878148 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 131;
Best Local Similarity 88.9%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 123 VESWILRNP 131

RESULT 4
Q8VLJ7_DEN4 PRELIMINARY; PRT; 131 AA.
AC Q8VLJ7;
DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2002, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Premembrane (Fragment).
GN Name=prM;
OS Dengue virus type 4.

```



```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF459620; AAL67822.1; -; Genomic RNA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR002535; Flavi_propep.
DR Pfam; PF01004; Flavi_M; 1.
DR Pfam; PF01570; Flavi_propep; 1.
FT NON_TER 1 1
FT NON_TER 131 131
SQ SEQUENCE 131 AA; 14975 MW; 10F1CD64F5364148 CRC64;

Query Match      88.5%; Score 46; DB 2; Length 131;
Best Local Similarity 88.9%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 123 VESWILRNP 131

RESULT 8
Q32PL2_DEN4
ID Q32PL2_DEN4 PRELIMINARY; PRT; 348 AA.
AC Q32PL2;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
FT NON_TER 1 1
FT NON_TER 131 131
SQ SEQUENCE 348 AA; 38409 MW; 8127734BF26F94E CRC64;

Query Match      88.5%; Score 46; DB 2; Length 348;
Best Local Similarity 88.9%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 9
Q32PL4_DEN4
```

```
ID Q32PL4_DEN4 PRELIMINARY; PRT; 348 AA.
AC Q32PL4;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Cairns;
RA Pyke A.T., Hanna J., Richards A., Taylor C.T., Morgan A.,
RA Humphreys J., Brookes D., Smith G.A.;
RT "Defining Dengue in the New Millennium.";
RL Arbovirus Res. Aust. 0:0-0(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY705987; AAW62440.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polypeptide.
FT NON_TER 1 1
FT NON_TER 348 348
SQ SEQUENCE 348 AA; 38423 MW; 4945CD4C6E5436D7 CRC64;

Query Match      88.5%; Score 46; DB 2; Length 348;
Best Local Similarity 88.9%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 10
Q32PL5_DEN4
ID Q32PL5_DEN4 PRELIMINARY; PRT; 348 AA.
AC Q32PL5;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polypeptide (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Cairns;
RA Pyke A.T., Hanna J., Richards A., Taylor C.T., Morgan A.,
RA Humphreys J., Brookes D., Smith G.A.;
RT "Defining Dengue in the New Millennium.";
RL Arbovirus Res. Aust. 0:0-0(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY705986; AAW62439.1; -; Genomic RNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
```

```

DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR00069; Flavi_M.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 348
SQ SEQUENCE 348 AA; 38409 MW; 8127734BF26F94EE CRC64;

Query Match      88.5%; Score 46; DB 2; Length 348;
Best Local Similarity 88.9%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 11
Q66432_DEN4
ID Q66432_DEN4 PRELIMINARY; PRT; 545 AA.
AC O66432;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Brazil 1982;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; U18425; AAB70676.1; -; Genomic_RNA.
DR HSPF; Q88653; IOKE.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0019058; P: viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR00069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT NON_TER 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
SQ SEQUENCE 545 AA; 59675 MW; 70E1CB02DCC7F33F CRC64;

Query Match      88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 12
Q66433_DEN4
ID Q66433_DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66433;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=El Salvador 1983;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
DR EMBL; U18426; AAB70677.1; -; Genomic_RNA.
DR HSPF; Q88653; IOKE.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0019058; P: viral infectious cycle; IEA.
DR InterPro; IPR011999; Flav_glyc_cen_dm.
DR InterPro; IPR00069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT NON_TER 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
SQ SEQUENCE 545 AA; 59675 MW; 70E1CB02DCC7F33F CRC64;

Query Match      88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 13
Q66435_DEN4
ID Q66435_DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66435;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Indonesia 1973;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----

```

```

CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U18429; AAB70679.1; -; Genomic_RNA.
DR HSP; Q88653; 10KE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
FT SEQUENCE 545 AA; 59667 MW; FBAD39FD161840CA CRC64;
Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15
RESULT 14
Q66436 DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66436;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DE 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Indonesia 1976;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U18429; AAB70680.1; -; Genomic_RNA.
DR HSP; Q88653; 10KE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
FT SEQUENCE 545 AA; 59667 MW; FBAD39FD161840CA CRC64;
Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15
RESULT 15
Q66437 DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66437;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DE 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Indonesia 1977;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U18430; AAB70681.1; -; Genomic_RNA.
DR HSP; Q88653; 10KE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
FT SEQUENCE 545 AA; 59645 MW; 9BAC3041F733AFB1 CRC64;
Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15
RESULT 16
Q66438 DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66438;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DE 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
FT SEQUENCE 545 AA; 59645 MW; 9BAC3041F733AFB1 CRC64;
Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15

```

```

OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Mexico 1984;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U18431; AAB70682.1; -; Genomic_RNA.
DR HSSP; Q88653; IOKE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flv_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Polyprotein.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON TER 1 1
FT NON TER 545 545
SQ SEQUENCE 545 AA; 59615 MW; 5636SD574E542122 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 17
Q66439 DEN4
ID Q66439 DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66439;
DT 01-NOV-1996, integrated into UniProtKB/TREMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=New Caledonia 1984;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U18432; AAB70683.1; -; Genomic_RNA.
DR HSSP; Q88653; IOKE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.

```

```

DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flv_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Polyprotein.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON TER 1 1
FT NON TER 545 545
SQ SEQUENCE 545 AA; 59675 MW; 70E1CB02DCC7F33F CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 18
Q66440 DEN4
ID Q66440 DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66440;
DT 01-NOV-1996, integrated into UniProtKB/TREMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Philippines 1956;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U18433; AAB70684.1; -; Genomic_RNA.
DR HSSP; Q88653; IOKE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flv_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dm.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
DR Polyprotein.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON TER 1 1
FT NON TER 545 545
SQ SEQUENCE 545 AA; 59662 MW; 431024CF932E77AC CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```





```

DR Pfam; PF00869; Flavi_glycoprot; 1.
KW Pfam; PF01004; Flavi_M; 1.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
SQ SEQUENCE 545 AA; 59631 MW; 3C31192DD2733B9DD CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 22
Q66444_DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66444;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polypeptide (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Tahiti 1979;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S.; Gubler D.J.; Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; U18438; AAB70688.1; -; Genomic_RNA.
DR HSP; Q88653; 10KE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polypeptide.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
SQ SEQUENCE 545 AA; 59673 MW; 0FB862CBD2C67063 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 24
Q66446_DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66446;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polypeptide (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Tahiti 1985;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S.; Gubler D.J.; Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; U18437; AAB70688.1; -; Genomic_RNA.
DR HSP; Q88653; 10KE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polypeptide.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
SQ SEQUENCE 545 AA; 59605 MW; 5D27B5A77AAA0FE2 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 23
Q66445_DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66445;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polypeptide (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Tahiti 1979;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S.; Gubler D.J.; Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; U18438; AAB70689.1; -; Genomic_RNA.
DR HSP; Q88653; 10KE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polypeptide.
FT CHAIN <1 50 membrane protein M.
FT CHAIN 51 >545 envelope protein E.
FT NON_TER 1
FT NON_TER 545
SQ SEQUENCE 545 AA; 59673 MW; 0FB862CBD2C67063 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VESWFLRNP 9
Db 7 VESWILRNP 15

```

```

-----
CC EMBL; U18439; AAB70690.1; -; Genomic_RNA.
DR HSSP; Q88653; LOKE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT NON_TER 51 >545 envelope protein E.
FT NON_TER 1 1
FT NON_TER 545 545
SQ SEQUENCE 545 AA; 59673 MW; 3545B1P5F87D11B4 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 25
Q66447_DEN4 PRELIMINARY; PRT; 545 AA.
AC Q66447;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Thailand 1963;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; U18440; AAB70691.1; -; Genomic_RNA.
DR HSSP; Q88653; LOKE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT NON_TER 51 >545 envelope protein E.
FT NON_TER 1 1
FT NON_TER 545 545
SQ SEQUENCE 545 AA; 59673 MW; 71AED00247DF7977 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 26
Q66448_DEN4 PRELIMINARY; PRT; 545 AA.
ID Q66448;
AC Q66448;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Thailand 1978;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; U18441; AAB70692.1; -; Genomic_RNA.
DR HSSP; Q88653; LOKE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT NON_TER 51 >545 envelope protein E.
FT NON_TER 1 1
FT NON_TER 545 545
SQ SEQUENCE 545 AA; 59623 MW; 71AED00247DF7977 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 27
Q66449_DEN4 PRELIMINARY; PRT; 545 AA.
ID Q66449;
AC Q66449;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.

```

```

SQ SEQUENCE 545 AA; 59666 MW; 588A14D46277DDA3 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 26
Q66448_DEN4 PRELIMINARY; PRT; 545 AA.
ID Q66448;
AC Q66448;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Thailand 1978;
RX MEDLINE=97437482; PubMed=9292015;
RA Lanciotti R.S., Gubler D.J., Trent D.W.;
RT "Molecular evolution and phylogeny of dengue-4 viruses.";
RL J. Gen. Virol. 78:2279-2284(1997).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
-----
EMBL; U18441; AAB70692.1; -; Genomic_RNA.
DR HSSP; Q88653; LOKE.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR011999; Flavi_glyc_cen_dm.
DR InterPro; IPR000069; Flavi_M.
DR InterPro; IPR000336; Flv_glyc_Ig-like.
DR InterPro; IPR011998; Vrl_glyc_cen_dim.
DR Pfam; PF02832; Flavi_glycop_C; 1.
DR Pfam; PF00869; Flavi_glycoprot; 1.
DR Pfam; PF01004; Flavi_M; 1.
KW Polyprotein.
FT CHAIN <1 50 membrane protein M.
FT NON_TER 51 >545 envelope protein E.
FT NON_TER 1 1
FT NON_TER 545 545
SQ SEQUENCE 545 AA; 59623 MW; 71AED00247DF7977 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 545;
Best Local Similarity 88.9%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 7 VESWILRNP 15

RESULT 27
Q66449_DEN4 PRELIMINARY; PRT; 545 AA.
ID Q66449;
AC Q66449;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE Polyprotein (Fragment).
OS Dengue virus type 4.

```

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus; Dengue virus group.  
 OX NCBI\_TaxID=11070;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Thailand 1984;  
 RX MEDLINE=97437482; PubMed=9292015;  
 RA Lanciotti R.S., Gubler D.J., Trent D.W.;  
 RT "Molecular evolution and phylogeny of dengue-4 viruses.";  
 RL J. Gen. Virol. 78:2279-2284(1997).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL: U18442; AAB70693.1; -; Genomic\_RNA.  
 DR HSSP; Q88653; 10KE.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0019028; C: viral capsid; IEA.  
 DR GO: GO:0019031; C: viral envelope; IEA.  
 DR GO: GO:0005198; F: structural molecule activity; IEA.  
 DR GO: GO:0019058; P: viral infectious cycle; IEA.  
 DR InterPro: IPR011999; Flav\_glyc\_cen\_dm.  
 DR InterPro: IPR000069; Flavi\_M.  
 DR InterPro: IPR000336; Flv\_glyc\_ig-like.  
 DR InterPro: IPR011998; Vrl\_glyc\_cen\_dm.  
 DR Pfam: PF02832; Flavi\_glycop\_C; 1.  
 DR Pfam: PF00869; Flavi\_glycoprot; 1.  
 DR Pfam: PF01004; Flavi\_M; 1.  
 KW Polyprotein.  
 FT CHAIN   
 FT CHAIN   
 FT NON\_TER   
 FT NON\_TER   
 SQ SEQUENCE 545 AA; 59623 MW; 5E74D2BD1B1B9A78 CRC64;  
 Query Match 88.5%; Score 46; DB 2; Length 545;  
 Best Local Similarity 88.9%; Pred. No. 9.5;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 VESWFLRNP 9  
 Db 7 VESWILRNP 15  
 RESULT 28  
 Q3ZPL3\_DEN4  
 ID Q3ZPL3\_DEN4 PRELIMINARY; PRT; 583 AA.  
 AC Q3ZPL3;  
 DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.  
 DT 20-DEC-2005, sequence version 2.  
 DT 07-FEB-2006, entry version 4.  
 DE Polyprotein (Fragment).  
 OS Dengue virus type 4.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus; Dengue virus group.  
 OX NCBI\_TaxID=11070;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ET00;  
 RA Pyke A.T., Hanna J., Richards A., Taylor C.T., Morgan A.,  
 RA Humphreys J., Brookes D., Smith G.A.;  
 RT "Defining Dengue in the New Millennium";  
 RL Arbovirus Res. Aust. 0:0-0(2005).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL: AY705988; AAW62441.2; -; Genomic RNA.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0019028; C: viral capsid; IEA.  
 DR GO: GO:0019031; C: viral envelope; IEA.  
 DR GO: GO:0019058; P: viral infectious cycle; IEA.  
 KW Polyprotein.

FT NON\_TER 1 1  
 FT NON\_TER 583 583  
 SQ SEQUENCE 583 AA; 63804 MW; 78FE1741AA8283DD CRC64;  
 Query Match 88.5%; Score 46; DB 2; Length 583;  
 Best Local Similarity 88.9%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 VESWFLRNP 9  
 Db 18 VESWILRNP 26  
 RESULT 29  
 Q68SA7\_DEN4  
 ID Q68SA7\_DEN4 PRELIMINARY; PRT; 590 AA.  
 AC Q68SA7;  
 DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.  
 DT 11-OCT-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 9.  
 DE Polyprotein (Fragment).  
 OS Dengue virus type 4.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Flavivirus; Dengue virus group.  
 OX NCBI\_TaxID=11070;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=No.14/Sri Lanka/Human;  
 RX PubMed=15302944; DOI=10.1099/vir.0.80120-0;  
 RA Mathenge E.G., Parquet M.D.C., Funakoshi Y., Houhara S., Wong P.F.,  
 RA Ichinose A., Hasebe F., Inoue S., Morita K.;  
 RT "Fusion PCR generated Japanese encephalitis virus/dengue 4 virus  
 RT chimera exhibits lack of neuroinvasiveness, attenuated neurovirulence,  
 RT and a dual-flavi immune response in mice.";  
 RL J. Gen. Virol. 85:2503-2513(2004).  
 RN (2)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=No.14/Sri Lanka/Human;  
 RA Mathenge E.G.M., Parquet M.D.C., Wong P.F., Ichinose A., Hasebe F.,  
 RA Inoue S., Morita K.;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL: AY559316; AAT37641.1; -; Genomic RNA.  
 DR GO: GO:0016021; C: integral to membrane; IEA.  
 DR GO: GO:0019028; C: viral capsid; IEA.  
 DR GO: GO:0019031; C: viral envelope; IEA.  
 DR GO: GO:0005198; F: structural molecule activity; IEA.  
 DR GO: GO:0019058; P: viral infectious cycle; IEA.  
 DR InterPro: IPR011999; Flav\_glyc\_cen\_dm.  
 DR InterPro: IPR000069; Flavi\_M.  
 DR InterPro: IPR002535; Flavi\_propep.  
 DR InterPro: IPR000336; Flv\_glyc\_ig-like.  
 DR InterPro: IPR011998; Vrl\_glyc\_cen\_dm.  
 DR Pfam: PF02832; Flavi\_glycop\_C; 1.  
 DR Pfam: PF00869; Flavi\_glycoprot; 1.  
 DR Pfam: PF01004; Flavi\_M; 1.  
 DR Pfam: PF01570; Flavi\_propep; 1.  
 KW Polyprotein.  
 FT NON\_TER 1 1  
 FT NON\_TER 590 590  
 SQ SEQUENCE 590 AA; 65142 MW; 748ELC79508D2A6E CRC64;  
 Query Match 88.5%; Score 46; DB 2; Length 590;  
 Best Local Similarity 88.9%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 VESWFLRNP 9  
 Db 123 VESWILRNP 131

```

RESULT 30
Q2YED3_DEN4
ID Q2YED3_DEN4 PRELIMINARY; PRT; 646 AA.
AC Q2YED3;
DT 20-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 20-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Polyprotein (Fragment).
OS Dengue virus type 4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus; Dengue virus group.
OX NCBI_TaxID=11070;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D4.CRA.1993;
RX PubMed=16282468; DOI=10.1128/JVI.79.23.14680-14687.2005;
RA Carrington C.V., Foster J.E., Pybus O.G., Bennett S.N., Holmes E.C.;
RT "Invasion and maintenance of dengue virus type 2 and type 4 in the
RT Americas."
RL J. Virol. 79:14680-14687(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D4.CRA.1993;
RA Carrington C.V.F., Foster J.E., Pybus O.G., Bennett S.N., Holmes E.C.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL; AY934757; AAY16569.1; -; Genomic_RNA.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 646 646
SQ SEQUENCE 646 AA; 70683 MW; E4A170ED7C6BC815 CRC64;

Query Match 88.5%; Score 46; DB 2; Length 646;
Best Local Similarity 88.9%; Pred.No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VESWFLRNP 9
Db 32 VESWILRNP 40
|||||
|

Search completed: August 31, 2006, 11:43:09
Job time : 140 secs

```